Effective Bandwidth Allocation for WiMAX Mesh Network

Hung-Chin JANG and Wei-Ching LIN
Department of Computer Science, National Chengchi University
Taipei, 11605, Taiwan, R.O.C.

ABSTRACT

The aim of this paper is to reduce the media access delay in a WiMAX mesh network. We observe that as the number of subscriber stations (SS) in a neighborhood increases, the processes of transmission opportunity (TO) competition and 3-way handshake are easy to fail. This may degrade transmission efficiency and increase packet transmission delay. Besides, the minislot allocation defined in the WiMAX mesh mode may cause many lower priority services reserve earlier minislots than that of higher priority services like rtPS. This may cause great negative impact on delay-sensitive traffic. In this paper, we design a QoS classifier to enqueue packets according to different QoS service classes, present a dynamic holdoff exponent mechanism to reduce control subframe delay, and propose a Neighborhood-Based Minislot Allocation (NBMA) mechanism to reduce data subframe delay. Simulations show that the proposed methodology outperforms that of IEEE 802.16 and Baye's DynExp in delay, jitter and throughput.

Keywords: WiMAX Mesh Network, Media Access, Delay, Transmission Opportunity, Bandwidth Allocation.

1. BACKGROUND

WiMAX mesh mode is composed of many interconnected mesh nodes, either subscriber stations (SS) or base stations (BS). These mesh nodes have to corporate by communicating with each other to route traffic through intermediate SSs to BS or gateway. In WiMAX mesh mode, packet scheduling can be either centralized scheduling or distributed scheduling. In centralized scheduling, SSs send their bandwidth requests to BS. BS will then allocate bandwidth to each SS according to both the available bandwidth and bandwidth demands from each SS. In distributed scheduling, bandwidth allocation is negotiated between the two communicating SSs. They have to spend time on determining the minislots to forward packets at each side. This is one of the main causes of packet delivery delay. Figure 1 shows the 802.16 mesh frame structure. Each frame consists of control sbframe for control message and data subframe for data packet. Each control subframe has 16 transmission opportunities (TO) with each TO being 7 OFDM symbols in length. Each data subframe is divided into multiple minislots (transmission data burst). The minislots are allocated through control messages exchange.

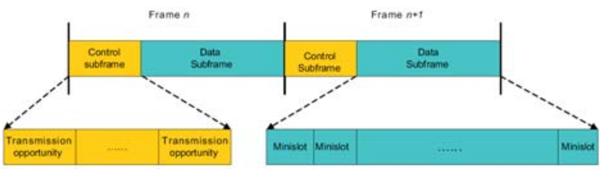

Figure 1: IEEE 802.16 Mesh Frame Structure

Bandwidth allocation is carried out in stages of TO contention and minislots allocation. SS competes TO for broadcasting MSH-DSCH (Mesh Distributed Schedule) message. MSH-DSCH is used to reserve minislots for data transmission between two SSs. The time interval between two consecutive TOs of one SS is defined as MSH-DSCH interval. One MSH-DSCH message consists of Request IE, Availability IE and Grant IE. These information elements (IE) are used in the process of 3-way handshake for minislots reservation. In the 3-way handshake, each SS broadcasts its requests, grants for neighbor SSs' requests, and available minislots to its neighbors. Each SS also broadcasts its confirmed minislots to its neighbors to avoid reservation conflict. In TO competition, each SS maintains values of holdoff exponent and mx. Holdoff exponent and mx are used to determine the length of holdoff time and the next transmission interval (nxmt interval), respectively. That is, SS will compete TOs in its nxmt interval and hold in its holdoff time. Holdoff exponent and mx should satisfy the following formulae.

$$HoldoffTime = 2^{(HoldoffExponent+4)}$$

$$2^{HoldoffExponent} \cdot mx < next\ transmission\ time \leq 2^{HoldoffExponent} \cdot (mx + 1)$$

In the literature, Bayer [1,2] analyzed the impact of holdoff exponent on MSH-DSCH interval. Bayer found that the larger the holdoff exponent, the longer the MSH-DSCH interval. The minislots reservation will thus be postponed resulting in deferring delay-sensitive traffic. Bayer proposed to use dynamic exponent to dynamically adjust the MSH-DSCH interval in order to reduce and increase TO competition frequencies of "inactive nodes" and "active nodes", respectively. Kuran [3] proposed to use "virtual node" concept to divide nodes with five service classes of UGS, ertPS, rtPS, nrtPS, and BE into five virtual nodes, each carries only one service class. These five virtual nodes individually send their bandwidth requests to BS. The BS then manages the admission control and bandwidth allocation according to different service classes. The advantage is that the limited bandwidth can be reserved for higher priority service classes in a highly congested network. Wongthavarawat [4] proposed to apply different packet scheduling mechanisms to four different service classes in an IEEE 802.16 network. Wongthavarawat suggested to allocate fixed bandwidth to UGS services to guarantee sufficient bandwidth, apply earliest deadline first (EDF) scheduling to rtPS services due to their real time constraints, apply weighted fair queue (WFQ) scheduling to nrtPS services due to their low real time but timeout constraints. BE services will be allocated bandwidth only if there is surplus bandwidth and the bandwidth is evenly distributed among these BE services.

2. METHODOLOGY

In TO competition, each SS executes a distributed election algorithm in its nxmt interval to decide whether it wins a candidate TO. The more SSs engaged in a TO competition, the less probability for each SS to win a TO. Equivalently, the longer time for each SS to win a TO due to participating more TO competitions. IEEE 802.16 standard defines holdoff time to prevent too many SSs from competing TOs all the time. Ideally, once a SS engages in a TO competition, most of the other SSs are in their holdoff time. In such case, the competition and delay are kept minimum. As to minislots allocation, once a SS wins a TO, it will be able to reserve the upcoming minislots no matter what kind of service the SS is going to transmit. It is very likely that many lower priority services book earlier minislots than those of higher priority services. This may cause severe packets delay.

In this paper, we propose three mechanisms to effectively solve the mentioned delay problem. In addition to those defined in the standard, we design a QoS classifier to enqueue packets according to different QoS service classes and use virtual connection id (virtual CID) of virtual node as QoS label. Those data come from the same virtual connection has the same QoS service type. We also propose a dynamic holdoff exponent mechanism to reduce control subframe delay by dynamically changing holdoff length according to the status of the neighbor SSs. Finally, we present a Neighborhood-Based Minislot Allocation (NBMA) mechanism to reduce data subframe delay by filtering out unimportant requests based on scoring.

2.1 Delay-Sensitive Request and Virtual Node

Different service classes have different QoS requirements on packet loss, delay and jitter. It is necessary to differentiate media access priorities accordingly to guarantee QoS. To be compatible with PMP mode, we define the same QoS service classes for mesh mode as those defined in PMP mode. That is, UGS, rtPS, nrtPS, and BE service classes. We employ multi-queue and virtual connection to separate distinct types of data streams such that delay-sensitive traffic can be served as early as possible and thus reduce packet delay. We adapt the virtual node concept of Kuran [3] to set up the virtual connections between the multi-queues of two communicating SS.

Requests are thereafter classified into UGS-request, rtPS-request, nrtPS-request, and BE-request. UGS-request and rtPS-request are delay-sensitive requests having more stringent constraints on delay requirement. On the other hand, nrtPS-request and BE-request are non-delay-sensitive requests.

2.2 Dynamic Holdoff Exponent

The length of holdoff time is determined by the holdoff exponent according to the formula: $HoldoffTime = 2^{(HoldoffExponent+4)}$. The smaller holdoff exponent, the smaller holdoff time and the higher competition frequency. Appropriate holdoff exponent will increase winning probability of each SS and thus reduce packet delay. We propose to dynamically adjust holdoff exponent such that those SSs having non-delay-sensitive and delay-sensitive packets are assigned longer and shorter holdoff time, respectively. Bayer [1,2] proposed to speed up 3-way handshake by reducing the holdoff exponents of Mesh-BS, sponsor node, and active node. However, problem occurs when the network traffic is congested and nodes will become active most of the time. In this case, the dynamic holdoff exponent mechanism unfortunately becomes "static" holdoff exponent mechanism. We modify Bayer's work by defining delay-sensitive (DS) and non-delay-sensitive (Non-DS) nodes to substitute for active and sponsor nodes, respectively. The proposed order of dynamic holdoff exponents becomes the following.

$$0 < \textit{mesh-BS} < \underline{DS} < \underline{\textit{Non-DS}} < \textit{non-active node} < 7$$

Ideally, once a SS engages in a TO competition, most of the other SSs are in their holdoff time. In other words, the first *nxmt* interval of each SS is exactly the holddoff time of the other neighbor SSs. In such case, it needs no competition but wins the TO at the first *nxmt* interval. Equivalently, the length of holdoff time of each SS is exactly long enough to have each neighbor SS send out one MSH-DSCH message. We can therefore have the following inequality hold.

$$2^{\exp+4} \geq \textit{number of neighbor}$$

Taking logarithm on both sides, we derive the minimum holdoff exponent as follows

$$\exp \quad \geq \quad \lceil \log(\textit{number of active neighbor}) \rceil - 4$$

Next, we define the dynamic exponent of each SS state as follows based on the above exponent order and inequality with lower bound set to 0.

- Mesh-BS: $\max\{0, \lceil \log(\textit{number of active neighbor}) \rceil - 5\}$
- Delay-sensitive: $\max\{0, \lceil \log(\textit{number of active neighbor}) \rceil - 4\}$
- Non-delay-sensitive: $\max\{0, \lceil \log(\textit{number of active neighbor}) \rceil - 3\}$
- In_active : 7

For a mesh BS, its state always remains Mesh-BS. For other SSs, state transfer is based on the changes of request IE types.

2.3 Request Scoring

As stated above, it is very likely that many lower priority services book earlier minislots than those of higher priority services like rtPS. This may cause severe packet delay problem. In this subsection we propose to assess the importance of each request by scoring. In the stage of minislot allocation, the available bandwidth will be allocated according to the ordering of importance. The importance of a request is defined as "the negative impact if the request is rejected". If a request is assessed as "large impact if rejected" then the request should be admitted with higher priority. That is, the larger the negative impact, the more important the request. We propose to assess the negative impact from packet loss, delay, and jitter perspectives.

2.3.1 Importance Factor of Packet Loss Rate: In a wireless network, the causes of packet loss may come from radio interference, network congestion, or TCP timeout. Packet loss caused by interference should be solved by the OFDM modulation or CRC mechanism of PHY layer. The PHY layer factors are beyond the scope of this paper. Both the network congestion and TCP timeout can be reflected by "the requested amount of minislots" and "available queue size". We therefore define the importance factor of packet loss rate (IM_PLR) as follows.

$$IM\_PLR = \frac{X}{[MaxQ - (CurQ - X)]}$$

Where X is the transmission amount during MSH-DSCH interval, MaxQ is the maximum queue size, and CurQ is the current used queue size.

2.3.2 Importance Factor of Delay: In IEEE 802.16 mesh network, each packet should be forwarded by a number of SSs before reaching the destination. Each forward needs one 3-way handshake. As one request gets rejected, it should wait for the next TO to send a new request to initiate a new 3-way handshake. All this process is definitely time consuming. Owing to both dynamic holdoff exponent and TO competition, the MSH-DSCH interval of requester may be greater than, equal to or less than that of granter. Figures 2-5 show numerous scenarios.

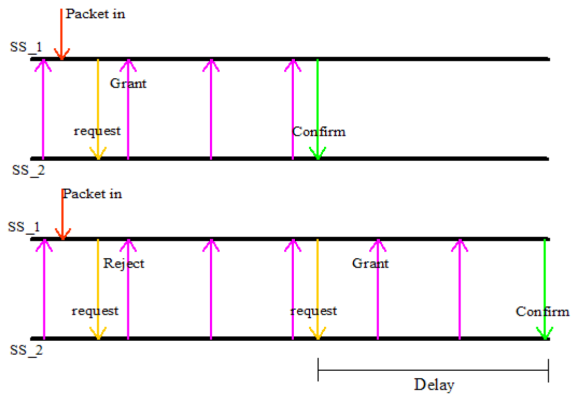

Figure 2: MSH-DSCH interval: SS_1 > SS_2

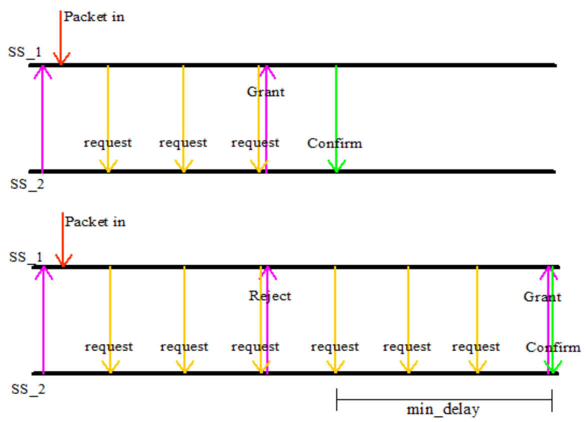

Figure 3: MSH-DSCH interval: SS_1 < SS_2
(Minimum Delay)

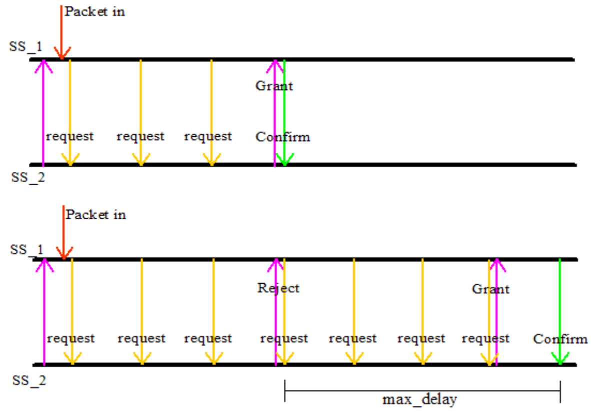

Figure 4: MSH-DSCH interval: SS_1 < SS_2
(Maximum Delay)

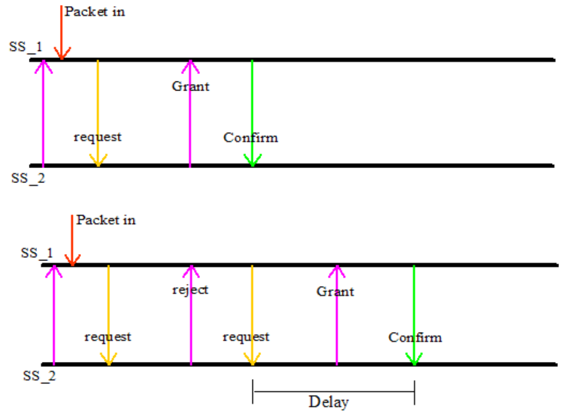

Figure 5: MSH-DSCH interval: SS_1 = SS_2

Each SS is able to estimate the "possible delay" by comparing its holdoff with that of the requester. In case of consecutive rejects, we accumulate all possible delays. We define the reject delay as the possible delay if the request is rejected.

$$reject\_delay = \sum_{k=0}^{reject\_count} MSH - DSCH\ length$$

Where the reject_count is the number of consecutive rejects and the MSH-DSCH length is the length of MSH-DSCH interval. The MSH-DSCH length depends on the relationship between the holdoff exponents of two SSs ($\exp_{ss}$) as follows:

$$\begin{cases} 2^{\exp_{ss\_1}+4} & , \quad if\ \exp_{ss\_1} > \exp_{ss\_2} \\ 2^{\exp_{ss\_2}+4} \pm 2^{\exp_{ss\_1}+4} & , \quad if\ \exp_{ss\_1} < \exp_{ss\_2} \\ 2^{\exp_{ss\_1}+4} & , \quad if\ \exp_{ss\_1} = \exp_{ss\_2} \end{cases}$$

The MaxLatencyPerHop is the tolerable delay on each hop and is defined as the ratio of the system maximum latency to the hop count to the BS.

$$MaxLatencyPerHop = \frac{SystemMaximumLatency}{Syn\_hop}$$

Where the "SystemMaximumLatency" is a system QoS parameter indicating the tolerable end-to-end latency for each QoS service class, and the "Syn_hop" is the number of hops to the BS. Finally, we define the importance factor of delay (IM_delay) to be the ratio of the "possible delay" to the MaxLatencyPerHop.

$$IM\_delay = \frac{reject\_delay}{MaxLatencyPerHop}$$

2.3.3 Importance Factor of Jitter: Jitter is the variation of delay due to intermittency of packet transmission. As a request is rejected, it may cause not only delay but also possible jitter. Similarly, the negative impact of jitter counts both single and consecutive rejects. The importance factor of jitter (IM_jitter) is defined as follows.

$$IM\_jitter = \frac{reject\_jitter}{SystemMaximumJitter}$$

Where the "SystemMaximumJitter" is the maximum tolerable jitter given by the system.

2.3.4 Scoring: Different service classes have different requirements on packet loss rate, delay and jitter. In assessing the importance of each service request, we use the following formula based on the information of Table1.

$$IM\_factor = w_1 * IM\_PLR + w_2 * IM\_delay + w_3 * IM\_jitter$$

IM_PLR, IM_delay and IM_jitter are normalized before use. w_1, w_2 and w_3 are weights of packet loss rate, delay and jitter, respectively. The values of these weights affect the importance factor distribution of each service class. The ISP is able to adjust these weights according to their pricing schema and traffic distribution.

Table1: Service Classes vs. QoS Sensitivity

	Packet loss	Delay	Jitter
UGS	Low	High	High
rtPS	Low	Medium	High
nrtPS	High	Medium	Low
BE	High	Low	Don't care

The afterwards grants of service requests should ensure that all those UGS and rtPS requests will be admitted earlier than those of nrtPS and BE.

2.4 Neighborhood-Based Minislot Allocation

In the bandwidth allocation defined in the 802.16 standard, it is very likely that many requests of lower priority being admitted earlier than that of higher priority. In this subsection we propose a Neighborhood-Based Minislot Allocation (NBMA) mechanism to solve this problem. The basic idea is that each SS should know the minimum threshold of importance factor within its neighborhood. Only those requests whose importance factors are greater than the minimum threshold will be allocated minislots. NBMA is used to prevent minislots from being allocated to those requests of lower priorities due to FCFS (First Come First Serve) while deferring those requests of higher priorities. The operation flow of NBMA is as follows: each request is assessed against its importance factor. The importance factor distribution of each SS will be made as an information element (IM-factor IE). The IM-factor IE is placed in a MSH-DSCH message and is then broadcast to its neighbor nodes. Each SS should also collect the importance factor distribution from the neighbor nodes and then update its own distribution accordingly. Each SS then determines a screening threshold of request acceptance according to the available bandwidth. This threshold will finally be used to determine whether a request will be granted or not.

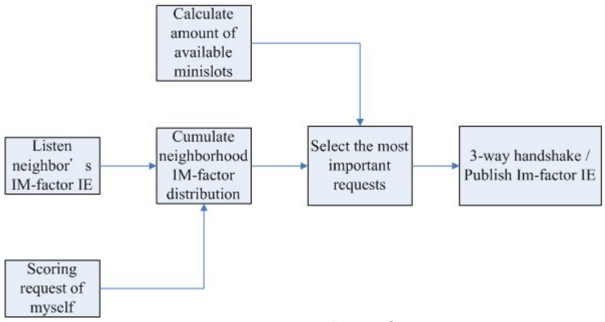

Figure 6: Operation Flow of NBMA

3. SIMULATIONS AND RESULTS

Simulations are performed on the NCTUns 4.0 simulator developed by Wang [5]. NCTUns 4.0 supports both WiMAX PMP and mesh modes. We modify the kernal by adding dynamic holdoff exponent and minislot allocation mechanisms. In the simulations, we dynamically adjust the holdoff exponent to check if it can reduce transmission delay effectively. We also verify that if SSs are able to differentiate bandwidth allocations according to the bandwidth requests from different service classes. The parameters setting is shown in Table 2.

Table 2: Parameters Setting of Simulations

OFDM modulation	64-QAM_3/4	108 bytes per OFDM symbol
Frame length code	4	Frame duration: 10 ms
MSH-CTRL-LEN	5	5 TOs per control subframe
MSH-DSCH-NUM	5	All TOs use distributed scheduling
Scheduling frame	2	Network control: Schedule control = 1:8

The network topology is shown in Figure 7 and all traffic is of UDP. Let the SS and SN generate UGS, rtPS, nrtPS and BE data flows at the average rate of 1.2Mbit/s. The performance metrics are delay, jitter and throughput. The proposed method, NBMA, is compared with both the original 802.16 and Bayer's DynExp.

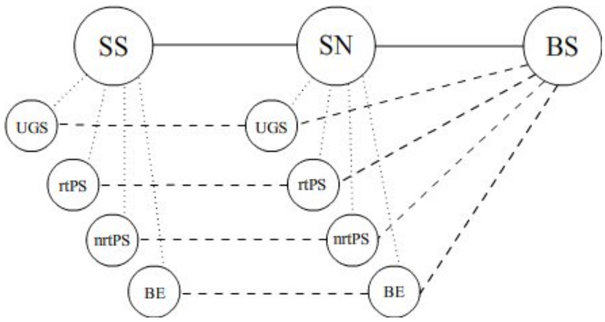

Figure 7: Network Topology of Simulations

Figure 8 shows the comparisons of delay. In a 2-hop environment, 802.16 mechanism has delay greater than 2 sec on average, while the NBMA has only 0.28~0.33 sec delay on average. NBMA improves 85% and 20% than that of 802.16 and DynExp in delay, respectively. This figure also indicates that holdoff exponent has great impact on delay. Figure 9 shows the comparisons of jitter. NBMA outperforms both 802.16 and DynExp by 12%~20% and 0.35%~7.5% improvement, respectively. Figure 10 shows the comparisons of throughput. NBMA has about the same performance as that of DynExp, but it outperforms 802.16 by 7.5%~8% improvement.

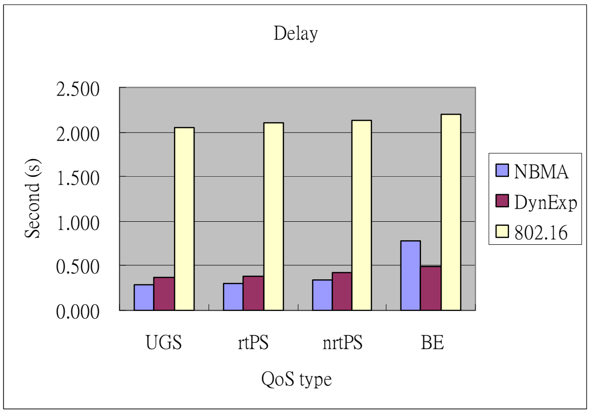

Figure 8: Comparisons of Delay

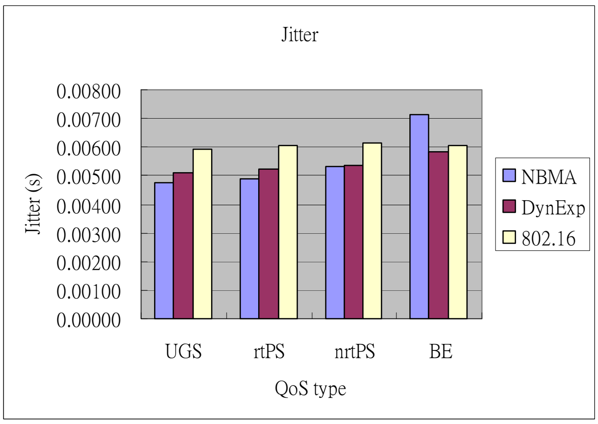

Figure 9: Comparisons of Jitter

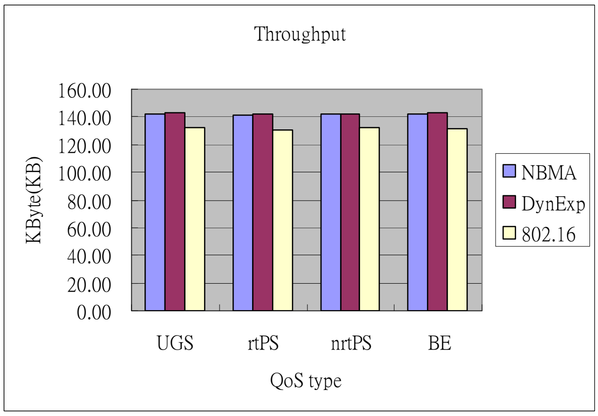

Figure 10: Comparisons of Throughput

4. CONCLUSIONS

To solve the problem of packet delay in the media access of WiMAX mesh mode, we design a QoS classifier to enqueue packets according to different QoS service classes, present a dynamic holdoff exponent mechanism to reduce control subframe delay, and propose a Neighborhood-Based Minislot Allocation (NBMA) mechanism to reduce data subframe delay. Simulations show that the proposed methodology outperforms that of IEEE 802.16 and Baye's DynExp in delay, jitter and throughput. In a 2-hop environment, the proposed NBMA improves 85% and 20% than that of 802.16 and DynExp in delay, respectively. NBMA has better performance in jitter than that of both 802.16 and DynExp by 12%~20% and 0.35%~7.5% improvement, respectively. Considering throughput, NBMA has about the same performance as that of DynExp, but it outperforms 802.16 by 7.5%~8% improvement.

5. REFERENCES

[1] N. Bayer, B. Xu, V. Rakocevic, and J. Habermann, “Improving the Performance of the distributed Scheduler in IEEE 802.16 Mesh Networks”, **Proceedings of Vehicular Technology Conference (IEEE VTC)**, Dublin, Ireland, April 2007.

[2] N. Bayer, B. Xu, V. Rakocevic, and J. Habermann, “Transmission Timing of Messages in IEEE 802.16 based Mesh Networks”, **Proceedings of IEEE/VDE European Wireless**, Athens, Greece, April, 2006.

[3] M.S. Kuran, B. Yilmaz, F. Alagoz, and T. Tugcu, “Quality of service in mesh mode IEEE 802.16 networks”, **14th International Conference on Software, Telecommunications and Computer Networks 2006**, (SoftCOM '06), Split-Dubrovnik, Croatia, 2006, pp. 107-111.

[4] K. Wongthavarawat, and A. Ganz, “Packet Scheduling for QoS support in IEEE 802.16 broadband wireless access systems”, **International Journal of Communication Systems**, Vol. 16, Issue 1, 2003, pp. 81-96.

[5] S.Y. Wang, C.L. Chou, C.H. Huang, C.C. Hwang, Z.M. Yang, C.C. Chiou, and C.C. Lin, "The Design and Implementation of the NCTUns 1.0 Network Simulator," **Computer Networks**, Vol. 42, Issue 2, June 2003, pp.175-197.