# QOS PROVISIONING FOR MOBILE AND FIXED WIRELESS REAL TIME MULTIMEDIA SERVICES

[1]Lakshmanan. M, [2]Noor Mohammed. V, [3]Nandakumar. S
[1,2&3] Assistant Professor, School of Electrical Sciences, VIT University, Vellore, India.
**Email:** < tmlakshmanan@gmail.com[1], noorb4u@gmail.com[2], nandha01@gmail.com[3]>

**Abstract:** In this paper, a new admission control policy and bandwidth reservation scheme for IEEE 802.16e networks has been proposed. Admission control policy and bandwidth reservation scheme are designed in such a way that desired quality of service is achieved. IEEE 802.16e standard did not specify any bandwidth reservation scheme that provides quality of service (QoS) support for real-time services, especially for handover real-time ones. The bandwidth reservation schemes, which were not designed for 802.16e system dedicatedly, could also be applied to it. But their performance is our main concern. Bandwidth reservation scheme is done based on probability of handover at mobile station and probability of call arrival at mobile station. And also admission control policy is designed for the same. Algorithm is designed to evaluate the performance of bandwidth reservation scheme. The simulation results show that not only minimizes the new service flow (SF) blocking rate and the handover SF dropping rate, but also reduces the access delay of new real-time SF.
**Keywords:** Admission control policy; Bandwidth reservation scheme; IEEE 802.16e; QoS; SF Blocking rate; SF Dropping rate

## 1. INTRODUCTION

IEEE 802.16e provides enhancements to support subscriber stations moving at vehicular speeds and thereby specifies a system for combined fixed and mobile broadband wireless access. Functions to support higher layer handover between base stations or sectors are specified. Operation is limited to licensed bands suitable for mobility below 6 GHz. Fixed IEEE 802.16 subscriber capabilities are not compromised.

This enables rapid worldwide deployment of innovative, cost-effective, and interoperable multivendor broadband wireless access products, facilitates competition in broadband access, encourages worldwide spectrum allocation, and accelerates the commercialization of broadband wireless access systems.

The initial 802.16 standard operates in the 10 to 66 GHz range. At these higher frequencies, IEEE 802.16 requires a direct line of sight between senders and receivers. This reduces multipath distortion, which occurs when broadcast signals not following a line of sight bounce off of large objects and end up out of synch, thereby scrambling the received transmission and decreasing bandwidth. Reducing multipath distortion, therefore, increases bandwidth. Theoretically IEEE 802.16 can provide single-channel data rates up to 75 Mbits per second on both the uplink and downlink. Providers could use multiple IEEE 802.16 channels for a single transmission to provide bandwidths of up to 350 Mbps.

## 2. QUALITY OF SERVICE

The Quality of Service (QoS) mechanism of the proposed baseline MAC consists of three basic components:

- Service flow QoS scheduling services
- Dynamic service establishment
- Two-phase activation model

The service flow QOS Scheduling services represent the data handling mechanisms supported by the Medium Access control (MAC) scheduler for data transport on a connection. Each connection is associated with a single data service type. There are five different data service types supported: Unsolicited Grant Service (UGS), Real-time Polling Service (rtPS), Extended rtPS, Non-real-time Polling Service (nrtPS), and Best Effort (BE).

### 2.1. THEORY OF OPERATIONS

The various protocol mechanisms described in this document may be used to support QoS for both uplink and downlink traffic through the SS and the BS.

The requirements for QoS include the following:

- A configuration and registration function for preconfiguring SS-based QoS service flows and traffic parameters.
- A signaling function for dynamically establishing QoS-enabled service flows and traffic parameters.
- Utilization of MAC scheduling and QoS traffic parameters for uplink service flows.
- Utilization of QoS traffic parameters for downlink service flows.

978-1-4244-4793-0/09/$25.00 © 2009 IEEE

- Grouping of service flow properties into named Service Classes, so upper-layer entities and external applications (at both the MS and BS) may request service flows with desired QoS parameters in a globally consistent way.

## 2.2. TWO PHASE ACTIVATION MODEL

The SF, which is the central concept of IEEE 802.16 MAC protocol, provides a mechanism for QoS management. Three basic types of SFs were defined in the standard [1]:

- Provisioned SF
- Admitted SF
- Active SF

A service flow may be provisioned but not immediately activated (sometimes called "deferred"). That is, the description of any such service flow contains an attribute that provisions but defers activation and admission [2].

This protocol supports a two-phase activation model that is often utilized in telephony applications. In the two-phase activation model, the resources for a "call" are first "admitted," and then once the end-to-end negotiation is completed (e.g., called party's gateway generates an "off-hook" event), the resources are "activated".

Active signal flow has resources omitted by the BS for transport of data packets. Only admitted or active SFs can be one-to-one mapped to connections and only active SFs may forward data packets. In the two-phase activation model [1], provisioned SF or dynamically created SF experiences admitted state firstly. Here, resources are reserved for this SF and admission control is performed. Then, the admitted SF is changed into active SF by dynamic service change (DSC) message exchange and the resources are activated, which completes the second stage of the model. DSA is dynamic service addition message. The block lines in Fig. 1 represent this process.

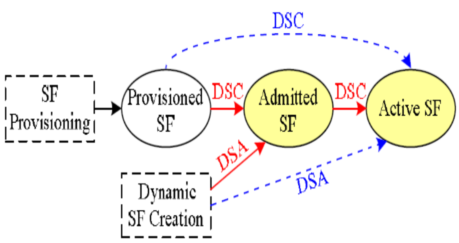


FIGURE 1 THE IMPLEMENTATION AND SKIP OF TWO PHASE ACTIVATION MODEL

But it is not always the case, especially for real-time services. A real-time SF may be provisioned and immediately activated. Similarly, a real-time SF may be created dynamically and instantly activated. In these cases, two-phase activation is skipped and the SF turns into active state directly without experiencing admitted state [1], When the free bandwidth is insufficient, the bandwidth requests of real-time SFs will be rejected.

## 2.3 BANDWIDTH ALLOCATION AND REQUEST MECHANISMS

During network entry and initialization every MS is assigned up to three dedicated connection identifiers for the purpose of sending and receiving control messages. These connection pairs are used to allow differentiated levels of QoS to be applied to the different connections carrying Medium access control layer (MAC) management traffic. Increasing (or decreasing) bandwidth requirements are necessary for all services except incompressible constant bit rate UGS connections. The needs of incompressible UGS connections do not change between connection establishment and termination. The requirements of compressible UGS connections, such as channelized T1, may increase or decrease depending on traffic. Demand assigned multiple access services are given resources on a demand assignment basis, as the need arises.

When an MS needs to ask for bandwidth on a connection with best effort scheduling service, it sends a message to the BS containing the immediate requirements of the demand assigned multiple access connection. QoS for the connection was established at connection establishment and is looked up by the BS.

## 2.4. BANDWIDTH RESERVATION SCHEME

Practically, WiMAX is the commercialization of the IEEE 802.16 standard. Because WiMAX is allowed to transmit at high power rates and use directional antennas to produce focused signals, it has a target range of up to 31 miles. But under the mobile situation, since the wireless links are vulnerable, MSs communicating directly with BSs likely will achieve a range of 5 to 6 miles [5]. In contrast, WiFi has a range of only several hundred feet and 3G cellular technology has a range of several thousand feet [4]. Compared with these wireless technologies, 802.16e network has a rather large coverage per BS, in which handover does not occur continually. In other words, the probability of handover is relatively low. In this case, the reserving bandwidth for both handover real-time traffics and potential new real-time traffics has been considered. Thus, not only the continuity of handover active real-time SF is guaranteed but also the access delay of new real time SF is reduced and the fairness of admitting new and handover real-time SFs is improved. The bandwidth is reserved for real-time traffic based on the probabilistic estimations of MS's handover and traffic arrival. Thus, a certain bandwidth range may be reserved for the real-time traffics belonging to different MSs simultaneously, not exclusively for a certain one. During the admission control process, the reserved bandwidth can be available to non-real-time traffic under the condition that the free bandwidth is insufficient and the reserved bandwidth is unused by real-time traffic.

Traffics issued from MSs are various. The traffic which is issued and the time when the traffic is issued are stochastic. Moreover, the transmission rate of each multimedia traffic flow changes with time in a random fashion [6]. Therefore, to design the bandwidth reservation scheme based on probability is reasonable. The bandwidth request and allocation mechanisms in the 802.16 standard specify that bandwidth is always requested on a connection basis and allocated on an MS basis [1][2]. For one thing, the BS collects the connections' bandwidth request information from the same MS, and grants the aggregate bandwidth to the MS. For another, the MS receives the grant and redistributes bandwidth among its connections, maintaining QoS and service level agreements [7]. Hence, the proposed DBQRS performs dynamic bandwidth reservation at BS taking MS as a unit. Aggregating traffics from several devices, each MS usually has relative steady traffics, not burst traffics [5]. Taking statistics about traffics issued from MS for a certain period of time, to determine the corresponding traffic models. In order to describe the proposed scheme, the parameters used in the paper are defined in Table.1.

| | |
|---|---|
| $M$ | The number of MSs quasi-reserved bandwidth by BS. |
| $N$ | The number of the real-time traffic types. |
| $P_m^{res}(t)$ | The probability of MS $m$ requiring BS to reserve bandwidth at time interval $t$. |
| $P_{m,n}^{res}(t)$ | The probability of the $n$-type traffic on MS $m$ requiring BS to reserve bandwidth at time interval $t$. |
| $P_{m,n}^{arr}(t)$ | The probability of the $n$-type traffic arriving at MS $m$ at time interval $t$. |
| $P_{m,n,s}^{arr}(t)$ | The probability of $s$ $n$-type SFs arriving at MS $m$ at time interval $t$. |
| $P_m^{ho}(t)$ | The probability of MS $m$ handover. |
| $\alpha_n$ | The weighting coefficient for the $n$-type real-time traffic. |
| $\beta_k$ | The weighting coefficient for MSs in different handover states. |
| $B_{m,n}$ | The reserved bandwidth required by the $n$-type traffic on MS $m$. |
| $B_m$ | The total reserved bandwidth required by MS $m$. |

TABLE 1: PARAMETER DEFINITIONS USED IN THE PAPER

Bandwidth reservation for the traffic of an MS is related not only to the traffic arrival, but also to the handover situation of the MS. From the multiplication theorem of probabilities,

Where
*P* (*bandwidth reservation for the traffic on the MS*)
= *P* (*handover of the MS*)
× *P* (*traffic arrival at the MS* | *handover of the MS*)
where $P(\bullet)$ and $P(\bullet|\bullet)$ denote the probability of the event occurring and conditional probability, respectively. Since the MS handover and traffic arrival at the MS are mutually independent, then
*P* (*traffic arrival at the MS* | *handover of the MS)*
*= P (traffic arrival at the MS)*
So that

$$P_{m,n}^{res}(t) = P_m^{ho}(t) \times P_{m,n}^{arr}(t).$$

The probability of MS m requiring BS to reserve bandwidth at time interval t is given by

$$P_m^{res}(t) = \sum_{n=1}^{N} \frac{s_{m,n}}{\sum_{n=1}^{N} s_{m,n}} P_{m,n}^{res}(t).$$

Substituting both the equation then resultant is

$$P_m^{res}(t) = P_m^{ho}(t) \times \sum_{n=1}^{N} \frac{s_{m,n}}{\sum_{n=1}^{N} s_{m,n}} P_{m,n}^{arr}(t) .$$

It is noted that the identical traffic model with different parameters results in the different values of traffic arrival probabilities. These model parameters change dynamically according to the real situation of traffics issued from each MS in the proposed scheme. Due to the different handover states as well as the diverse QoS requirements of traffics, to weight these probability values using the different weighting coefficients. Mobile users are more sensitive to terminating an ongoing real-time service than blocking a new one, handover real-time SFs are usually given higher priority over the new SFs. From the viewpoint of BS, three handover states has been considered. $\beta k$ is defined on the set $\{\beta in, \beta stay, \beta out\}$ with $\beta in > \beta stay > \beta out$ and each $\beta k$ corresponds to one handover state.

- $\beta in$ is used for the MSs that have active real-time SFs and are likely to immigrate into the BS's coverage area from the neighbor cells.
- $\beta stay$ is applied to the MSs that hardly move, stay in the BS's coverage area and send real-time traffics prospectively. In this case, probability of MS handover is always small, even approaching zero. In order to reserve bandwidth for the real-

time SFs on these local MSs, let probability of MS handover be 1.

- *βout* is for the MSs that are likely to move out of the serving BS's coverage area. In this case, probability of MS handover is replaced by its complement, (1 – Probability of MS handover)

So accordingly the above equation is transformed into

$$P_m^{res}(t) = \beta_k P_m^{ho}(t) \times \sum_{n=1}^{N} \frac{s_{m,n}}{\sum_{n=1}^{N} s_{m,n}} \alpha_n P_{m,n}^{arr}(t).$$

As from above equation, the larger $\alpha_n$ and $\beta_k$ are, the larger the probability of MS m requiring BS to reserve bandwidth at time interval t is, which means BS reserving bandwidth for MS *m* with greater probability. Thus, traffic obtains its desired bandwidth with greater possibility during the admission control process. If the probability of MS m requiring BS to reserve bandwidth at time interval t = 1, the bandwidth will be fully reserved for MS *m*, which is the same as conventional bandwidth reservation. Full-bandwidth reservation induces low bandwidth utilization. Therefore, taking the tradeoff between the efficiency of bandwidth reservation and the bandwidth utilization into account, $\alpha_n$ and $\beta_k$ should be designed.

The proposed scheme allows a certain bandwidth range to be quasi-reserved for the traffics belonging to different MSs simultaneously [8]. Thus, the below should be satisfied in the certain bandwidth range during time interval *t*; otherwise, the quasi-reservation will be performed in the other available bandwidth ranges.

$$\sum_{m=1}^{Mr} P_m^{res}(t) \le 1, \qquad Mr \le M$$

where *Mr* is the number of MSs with quasi-reserved bandwidth in the same bandwidth range during time interval *t*. With regard to the reserved bandwidth required by MS *m*, *Bm*, then

$$B_m = \sum_{n=1}^{N} B_{m,n} = \sum_{n=1}^{N} b_{m,n} \times s_{m,n},$$

where *bm,n* is the required bandwidth per *n*-type SF on MS *m*. Thus, the mean reserved bandwidth for MS *m* is estimated as

$$\overline{B}_m = B_m \times P_m^{res}(t).$$

In the real system, the serving BS monitors the position of MS, the velocity and direction of MS movement. Based on the information, the serving BS estimates probability of handover for MS *m*. Meanwhile, the serving BS calculates , the probability of n-type traffic arriving at MS m at time interval t and $B_m$ according to traffic models. It is noted that bandwidth is quasi-reserved for the handover MS at both target BS and serving BS using different handover probabilities, probability of MS handover and (1 – probability of MS handover), respectively. The BS performs dynamic bandwidth quasi-reservation for MSs according to the probability of MS m requiring BS to reserve bandwidth at time interval t and total reserved bandwidth required by MS m. The following table 2 shows the traffic models used in 802.16e networks.

| Traffic Type | Real-time traffic | | |
|---|---|---|---|
| | UGS | RT-VR | ERT-VR |
| Application | Voice | Video | VoIP with silence suppression |
| Inter-packet Time | Constant mean: 100 ms | Gamma mean: 35-90 ms std dev: 10 ms | Exponential mean: 50-100 ms |
| Packet Size | Constant mean: 64 bytes | Gamma mean: 64 bytes std dev: 3 bytes | Exponential mean: 32 bytes |
| Traffic Type | Non-real-time traffic | | |
| | NRT-VR | BE | |
| Application | File transfer | E-mail | |
| Inter-packet Time | Geometric mean: 200 ms | Exponential mean: 3600 ms | |
| Packet Size | Pareto mean: 1024 bytes shape parameter: 2 | Exponential mean: 500 bytes | |

TABLE 2 TRAFFIC MODELS

UGS, RT-VR and ERT-VR were considered to be equal as real time services and NRT-VR and BE to be equal as non real time services. So therefore real time services are prioritized over non real time services. And mobile users are more sensitive to terminating an ongoing real-time service than blocking a new one, handover real-time SFs are usually given higher priority over the new SFs. When bandwidth is reserved traffic obtains its desired bandwidth with greater possibility during the admission control process. Full bandwidth shouldn't be reserved as it will induce low bandwidth utilization.

When a handover MS immigrates into the BS's coverage area, the MS can gain the required bandwidth for its ongoing SFs immediately. At the same time, the BS quasi-reserves bandwidth for local MS, so that once a new real-time SF belonging to the MS requests bandwidth, it can obtain desired bandwidth more easily. Thereby, the fairness of getting bandwidth between handover and local real-time SFs is improved.

### 2.5. ADMISSION CONTROL POLICY

The purpose of admission control is the admission of only as many new users in the system that can maintain the quality of service promised to users already admitted. Admission control is implemented whenever a handover

MS with active SFs enters the BS's coverage or a local MS issues new SFs.

When a real-time SF of MS requests bandwidth, BS releases the previous reserved bandwidth for MS and recalculates free bandwidth. Here, MS may be a handover or local one. If free bandwidth meets the requirement of SF, SF is admitted; otherwise it is rejected. As for non-real-time SF, based on the quasi-reservation concept, it can occupy the reserved bandwidth temporarily when the free bandwidth is insufficient and real time SFs are not using the reserved bandwidth. Once real-time SFs arrive, non-real-time SFs must release the reserved bandwidth occupied by them for real-time SFs. These operations can be realized by BS initiating DSC message to change the state of non-real-time SFs to be inactive. Then, after a certain back off time, non-real-time SFs re-request bandwidth or piggyback bandwidth request with data.

## 3. PERFORMANCE EVALUATION

In order to evaluate the performance of proposed Bandwidth Reservation Scheme, two phase activation model is taken into comparison. Simulation environment is composed of a cell having 100 channels for signal flow allocation and having 2 handover MSs to neighboring cells and 4 handover MSs from neighboring cells in a period of 3 minutes. Average call duration is also considered to be 3 minutes. Two phase activation model doesn't reserve the bandwidth even for the handover MS whereas the proposed bandwidth reservation scheme reserves the bandwidths for the SFs.

Figs. 2 and 3 illustrate the effects of diverse schemes on the new real-time SFs in terms of new SF blocking rate (NSBR) and access delay, respectively. Here, access delay is defined as the time between the MS sending bandwidth request message for the SF and the SF obtaining its desired bandwidth. Due to reserving bandwidth for new real-time SFs in the BRS, these SFs can obtain required bandwidth as soon as possible; thereby the NSBR and the access delay are reduced. In TPAS it causes that the most real-time SFs cannot obtain desired bandwidth, especially when more SFs arrive. So that both NSBR and access delay are higher than those of the BRS. In Fig. 4, the handover SF dropping rate (HSDR) of BRS is lower than that of TPAS. In the TPAS, bandwidth isn't reserved for handover SF dedicatedly. When the overall traffic load is heavy, the bandwidth requirement of handover SF is satisfied difficultly. In the BRS, though bandwidth is reserved for each MS with real-time SFs, the reserved bandwidth can be assigned to non-real-time SFs. When real-time SFs arrive, BS can control these non-real-time SFs to release the reserved bandwidth.

In the proposed bandwidth reservation scheme if there is no bandwidth for arriving signal flows of the real time services then the non real time services occupying the bandwidth is dismissed and the bandwidth is allocated for the arrived signal flow. So new signal flow blocking rate is reduced when compared to two phase activation model shown in the fig.2.

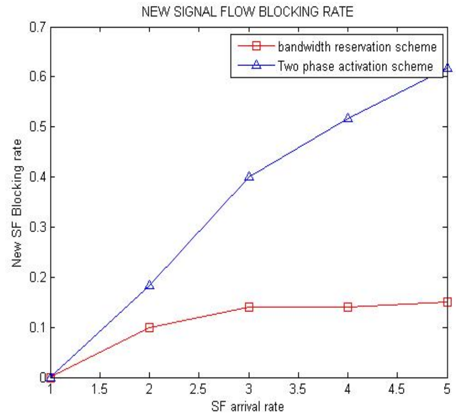


FIGURE.2 NEW SIGNAL FLOW BLOCKING RATE

In two phase activation model the calls are queued up when the signal flow arrival rate increases. So as a result of that delay factor increases as the signal flow arrival rate increases. In the proposed bandwidth reservation scheme since the some of bandwidth are reserved, the delay factor is reduced. So when the reserved bandwidth is free, access delay won't be there. Therefore access delay is less in the proposed bandwidth reservation scheme when compared to two phase activation model and it is shown graphically in fig.3.

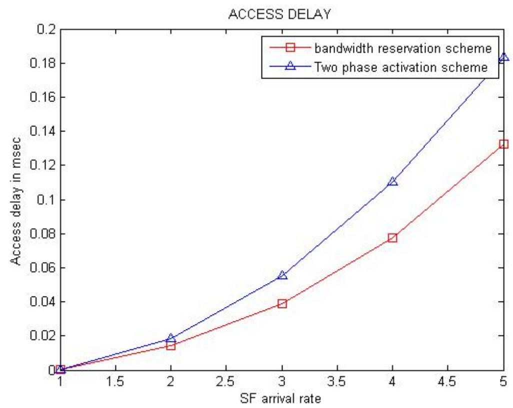


FIGURE 3 ACCESS DELAY

In two phase activation model when handover occurs if there is no free bandwidth available then the handover signal flow is rejected. But in bandwidth reservation scheme if there is no free bandwidth available then the bandwidth occupied by non real time services are freed and the bandwidth is made free for the handover signal flows and the arriving signal flow. Therefore signal flow blocking rate is minimized in case of bandwidth reservation scheme and it is shown graphically in fig 4.

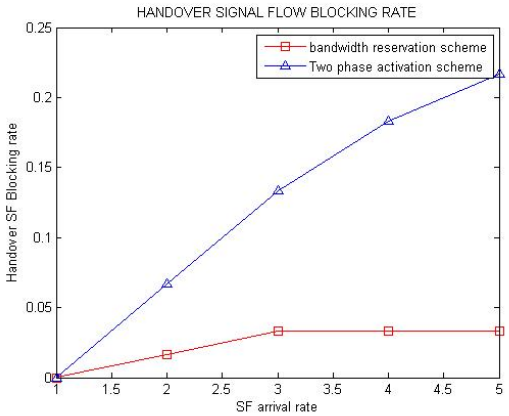


FIGURE 4 HANDOVER SIGNAL FLOW BLOCKING RATE

## 4. CONCLUSION

In this paper, considering that no appropriate bandwidth reservation scheme was specified for real-time services in IEEE 802.16e standard, the BRS and the corresponding admission control policy has been designed. According to the handover probability, the traffic arrival probability and the desired bandwidth of traffics, bandwidth is quasi-reserved for mobile and fixed wireless real-time multimedia services in 802.16e networks. Moreover, the reserved bandwidth can be adjusted dynamically. In addition, when the reserved bandwidth is idle, it can be occupied by non-real-time service; when real-time service arrives, the reserved bandwidth can be released. The proposed scheme not only provides QoS guarantee for real-time services, but also ensures the fairness of admitting handover and new real-time services. The simulation results demonstrate that the proposed scheme achieves low NSBR and HSDR, low access delay for new real-time service

## REFERENCES

[1]. IEEE Std 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air Interface for Fixed Broadband Wireless Access Systems," on Oct. 1, 2004.

[2]. IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor1-2005, "IEEE Standard for 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," on Feb. 28, 2006.

[3]. Dusit Niyato and Ekram Hossain, "Call Admission Control for QoS Provisioning in 4G Wireless Networks: Issues and Approaches," *IEEE Network*, vol. 19, no. 5, pp. 5-11, Sept./Oct. 2005.

[4]. Steven J. Vaughan-Nichols, "Achieving Wireless Broadband with WiMAX" *Computer*, vol. 37, no. 6, pp. 10 – 13, Jun. 2004.

[5]. Zakhia Abichar, Yanlin Peng, and J. Morris Chang, "WiMAX: The Emergence of Wireless Broadband," *IT Professional*, vol. 8, no. 4, pp. 44 - 48, July – Aug. 2006.

[6]. Jun Ye, Xuemin (Sherman) Shen, and Jon W. Mark, "Call Admission Control in Wideband CDMA Cellular Networks by Using Fuzzy Logic," *IEEE Trans. on Mobile Computing*, vol. 4, no. 2, pp. 129 – 141, Mar./Apr. 2005.

[7]. Qing Huang and G. S. Kuo, "IEEE 802.16 Based QoS Architecture for Future Fixed Broadband Wireless Access Systems," *Proc. of the 12th Meeting of WWRF*, Toronto, Canada, on Nov. 4 – 5, 2004.

[8]. G. S. Kuo, Po-Chang Ko and Min-Lian Kuo, "A Probabilistic Bandwid Estimation and Semi-Reservation Scheme for Flow-Oriented Multimedia Wireless Networks," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 135 – 141, Feb. 2001.

[9]. Xiang Chen, Bin Li, and Yuguang Fang, "A Dynamic Multiple-Threshold Bandwidth Reservation (DMTBR) Scheme for QoS Provisioning in Multimedia Wireless Networks," *IEEE Trans. on Wireless Communications*, vol. 4, no. 2, pp. 583 – 592, Mar. 2005.

[10]. Lei Huang, Sunil Kumar and C.-C. Jay Kuo, "Adaptive Resource Allocation for Multimedia QoS Management in Wireless Networks," *IEEE Trans. on Vehicular Technology*, vol. 53, no. 2, pp. 547 – 558, Mar. 2004.