A new QoS Guarantee Strategy in IEEE802.16 MAC Layer

Guannan Qu[1], Zhiyi Fang[1], Wei Cao[2], Youlai Yan[1], Ansheng Ma[1]
[1]*College of Computer Science and Technology, Jilin University, P.R.CHINA*
[2]*College of Information Technology, Liaodong University, P.R.CHINA*
qgn_0316@163.com, zyfang@public.cc.jl.cn

Abstract

IEEE802.16, a standard for wireless MAN broadband access technology, provides a flexible QoS mechanism in MAC layer for wild variety classes of services. However, it fails to define specific admission control strategy and scheduling algorithm, therefore, those aspects about QoS guarantee mechanisms are becoming the research hotspot in recent years. In this paper, focusing on the features of services defined in IEEE802.16, including UGS, ertPS, rtPS, nrtPS and BE, We propose a new admission control strategy based on the minimum reserved bandwidth and design a scheduling algorithm to ensure the implement of this strategy. Finally, we simulate the admission strategy and the algorithm on NS2 platform to validate their effectiveness.

Keywords: Wireless networks, QoS, IEEE 802.16, MAC, Admission control

1. Introduction

At present, fiber, as the main medium for the backbone network, has basically met the bandwidth request. However, the main bottleneck lies on the "last kilometer" problem, namely the bottleneck exists between the user and the backbone networks. The broadband wireless access system has provided an effective way to solve this problem, and IEEE802.16 standard offers a global unified standard for the broadband wireless access. Unfortunately, compared with the wire transmission, there are more constraints in the wireless transmission, such as the limited wireless bandwidth, the susceptible transmission, the low link reliability, the equipment battery energy problem and mobile handoff problem. Therefore, it is necessary to provide all kinds of services in the network transmission with essential QoS support.

The prominent characteristic of the IEEE802.16 standard is to provide a flexible QoS mechanism in the MAC layer, which is mainly composed of two parts: one part is the management mechanism such as the division of the scheduling service type, the dynamic service flow management, the polling, the signaling architecture of broadband request/distribution. They have been detailedly defined in the IEEE802.16 standard; the other one is the corresponding guarantee mechanisms, including admission control, scheduling algorithm, buffer area management and so on, which have not been carried on the definition and the elaboration in the standard. Apparently, a good admission control mechanism and an efficient scheduling algorithm not only can optimize the limited wireless resources, but also improve the wireless network broadband utilization, which largely satisfy the users' QoS requirements. We can see that more users can get the better service. As the result, these guarantee mechanisms are becoming the research hotspot in recent years.

QoS in 802.16 is supported by allocating each connection between the SS and the BS (called a service flow in 802.16 terminologies) to a specific QoS class. In 802.16e, there are 5 QoS classes: unsolicited grant service (UGS), real-time polling service (rtPS), non-real-time polling service (nrtPS) and best effort (BE) service. A new service called extended real-time polling service (ertPS) is added to IEEE208.16e standard. In this paper, we propose a new QoS admission control strategy according to the characteristics of these services, and design an algorithm to support the strategy. With this strategy, apart from BE business, all businesses will be reserved minimum bandwidth, which make each business can get service even when the network load is heavy.

2. Admission control strategy based on the minimum reserved bandwidth

Under the admission control strategy based on the shared bandwidth, when the UGS arrives at a high rate, lots of low-priority services are possible to be blocked, which results in unfair business among services [1-4], therefore, we propose a new admission control strategy based on the minimum reserved bandwidth in this paper. Considering fairness, it is designed to guarantee a part of low-priority services admission from starving to death, even when a large number of high-priority

978-1-4244-2020-9/08/$25.00 ©2008 IEEE

services fill the network, meanwhile it can help to make some high-priority services be admitted preferentially.

2.1 The basic thought

The basic thought of the admission control strategy based on the minimum reserved bandwidth is described as follow: Above all, we divide the whole bandwidth into several parts: one of them is assigned to be a public reserved bandwidth, and others are distributed to each class of services, just be illustrated as fig 2.1. When a new service ask for admission, its corresponding reserved bandwidth will be estimated whether to meet the demand. It will obtain bandwidth directly if its corresponding reserved bandwidth can satisfy its demand. Otherwise, the options will be executed as follow: For ertPS, rtPS and nrtPS, the services which have been admitted their reserved bandwidth are degraded under the precondition of ensuring their own QoS. That is, the transmission rate of these services will be reduced to its minimum, which helps to release some bandwidth for the new service requirement. If the corresponding reserved bandwidth after releasing can satisfy the requirement, the new service is able to admit directly, otherwise, it will use the pubic reserved bandwidth; For UGS, as its fixed-size data packets, the new service can be assigned the public reserved bandwidth without degrading; For BE, as it is a kind of best effort service, it occupies public reserved directly. It will be reduced QoS until zero to meet the shortage of public reserved bandwidth.

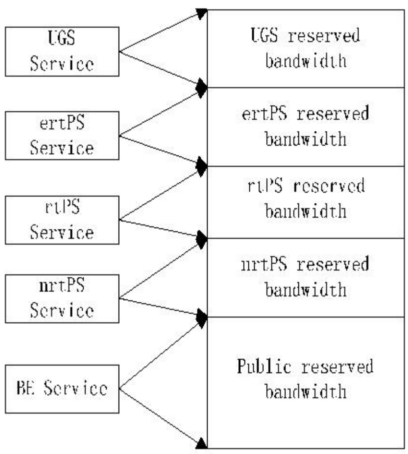

Fig 2.1 Reserved Bandwidth

2.2 Process of UGS admission control

UGS services for Real-time data streams, comprising fixed-size data packets issued at periodic intervals, therefore the maximum of transmission rate is equal to the minimum. However, for ertPS, rtPS, nrtPS and BE service, their service flow are unfixed, thereby their connection transmission rate are between the maximum and the minimum.

When a new UGS service ask for admission, if its corresponding reserved bandwidth remainder bigger than the minimal requirement (rate of the service flow), that if $C_i - C_{i\_used} \geq r_{\min}$ is true, it can be admitted; Else the public reserved bandwidth remainder will be estimated whether it is able to satisfy the requirement, that if $C_{res} - C_{res\_used} \geq r_{\min}$ is true, it can be admitted; If not, BE service which has been admitted to the public reserved bandwidth should be reduce until the remainder can satisfy the requirement, that if $\sum_{j=1}^{J} C_{be\_reduce}(j) \geq r_{\min}$ is true; If all BE service has reduce to zero, the remainder is still not enough, the admission requirement will be rejected.

The denotations are defined as follow:

C_i: Reserved bandwidth for the i class service.

C_{res}: Public reserved bandwidth.

C_{res_used}: The part of public reserved bandwidth which has been used.

C_{i_used}: The part of reserved bandwidth of the i class service which has been used

C_{be_reduce}: The bandwidth which be released after one BE service being degraded.

$\sum_{j=1}^{J} C_{be\_reduce}$: Summation of the bandwidth which be released after j BE service being degraded.

C_i can be evaluated by historical service flux loaded in network, and it is generally as a constant. C_{i_used} is dynamic real-time flux in network.

2.3 Process of ertPS, rtPS and nrtPS admission control

For ertPS、rtPS、nrtPS, the process of admission control is as follow: When a new service asks for admission, firstly it will estimate whether its corresponding reserved bandwidth remainder bigger than its requiring minimal rate, that if $C_i - C_{i\_used} \geq r_{\min}$ is true, it can be admitted. Else the services flow which has been admitted should be degraded under the precondition of ensuring their QoS. After this option, it will estimate whether it can satisfy the requirement, that if $\sum_{j=1}^{J} C_{i\_reduce}(j) \geq r_{\min}$ is true, it can be admitted. If the remainder is still not enough to meet the requirement after degrading the admitted service, the public remainder bandwidth will be estimated, that if $C_{res} - C_{res\_used} \geq r_{\min}$ is true, it can be admitted; Else BE , admitted services in the public reserved area, should be degraded until the remainder can satisfy the requirement, that if $\sum_{j=1}^{J} C_{be\_reduce}(j) \geq r_{\min}$ is true, it will be admitted. However, if there is no enough bandwidth even when the BE's $r_{\min}$ has been reduced to 0, it will be rejected to admit.

The reserved bandwidth for ertPS, rtPS and nrtPS can guarantee their OoS to some extent, even when UGS overload the network. The public reserved bandwidth, however, can be distributed to UGS, erPS, rtPS and nrtPS to meet the shortage of their own reserved bandwidth. UGS can use public reserved bandwidth preferentially, which holds the high priority of the UGS service. Despite scarce guarantee for BE service, as a class of data streams for which no minimum service level is required, it can be tolerated.

3. A multi-level scheduling algorithm according to the service classes

There are three well known scheduling algorithms, including Channel State Dependent Packet Scheduling (CSDPS) algorithm, Channel State Dependent Packet Scheduling + Class-Based Queuing(CSDPS+CBQ) algorithm and Idealized Wireless Fair Queuing (IWFQ) algorithm. We describe briefly as follow:

Channel State Dependent Packet Scheduling (CSDPS) algorithm is a kind of packet scheduling strategy based on device diver. As consideration of channel states, it can help to provide higher data throughput and higher channel usage. However, it is short of bandwidth guarantee mechanism for SS, which lead to less chance of obtaining service than average under bad connection.

In order to solve the unfairness of the CSDPS algorithm wireless bandwidth share, C Fragouli propose a strengthen algorithm, which is combined with Class Based Queuing [5, 6], that is CSDPS+CBQ algorithm. In this scheme, data stream is divided into packets, and each packet is assigned a certain bandwidth. CSDPS module is designed to deal with change of wireless connection, and CBQ module is used to provide a fair wireless connection sharing mechanism. This algorithm can achieve a fair share of the wireless channel under the condition of maintaining high throughput, but it does not defined a clear mechanism to compensate the user who has lost the opportunity of sharing service due to link trouble.

IWFQ is defined according to the reference model, which is a scheduling model of error-free and good state WFQ system. It ensures the fairness in bandwidth allocation, and carries out a limited bandwidth compensation for the system when the link gets troubles due to delay flow and lead flow.

Unfortunately, these algorithms are not adjusted to IEEE802.16 standard MAC layer, and direct usage of these algorithms can not satisfy the QoS requirement of all classes of service [7-10], as well as it can not ensure the implement of the admission control strategy based on the minimum reserved bandwidth. Therefore, combining the advantages and disadvantages of various algorithms with the characteristic of service classes in MAC layer, we discuss a multi-level scheduling algorithm.

3.1 Thought of the multi-level scheduling algorithm

Two main aspects are considered in designing multi-level scheduling algorithms: on one hand, more real-time requirement needs more services; on the other hand, the scheduling fairness should be ensured that a large number of low-priority operations will not be "starved to death."

According to these requirements, we adopt IWFQ algorithm in the first level schedule, for the IWFQ algorithm can ensure fairness of the wireless schedule. Considering the specific wireless transmission, we set different value to B, the parameter of the algorithm, to ensure that services could get corresponding compensation after resuming from link failure, that is, less bandwidth compensation will be for delay-sensitive service, as well as more bandwidth compensation will be for delay-insensitive service.

For different services, different scheduling algorithms should be adopted in the second level schedule. For UGS, ertPS, as they service for the connection real-timely and periodically, we adopt periodical schedule for them; for rtPS, nrtPS and BE we use the Weighted Round Robin algorithm (WRR). We set weight according to priority: rtPS will be set larger weight as its high real-time requirement to ensure more round times, and nrtPS and BEof will be set lower weight for their low or no real-time requirement. As the admission control strategy proposed in above section, we have reserved the minimum bandwidth for rtPS, nrtPS service, which ensure QoS of these service be satisfied without "starving to death", even when a large number of UGS and ertPS fill the network.

3.2 Architecture of the multi-level scheduling algorithm

The architecture of the multi-level scheduling algorithm is illustrated as fig 3.1.

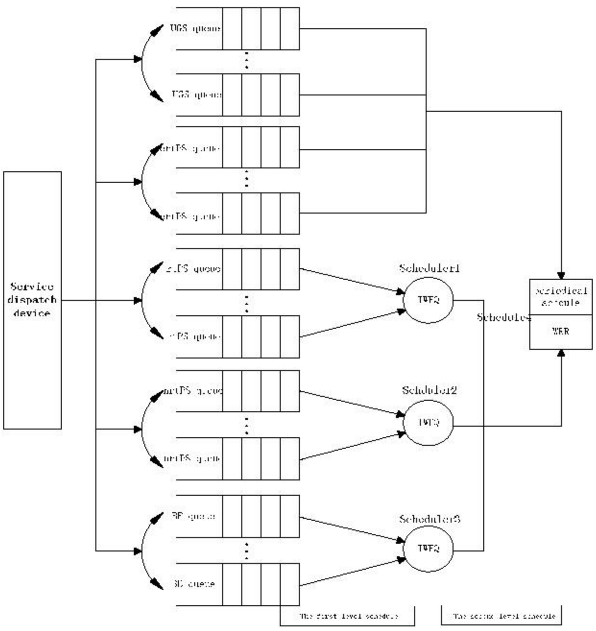

Fig 3.1 Architecture of the multi-level scheduling algorithm

Adopting IWFQ algorithm, the scheduler 1 takes charge of scheduling rtPS services. It maintains a queue for each rtPS service flow. When the number i packet of the rtPS service flow arrived, it will be signed the start time $S_{i\_rtps,n}$ and the final time $f_{i\_rtps,n}$. The relationship between the start time and the final time is as follow:

$$S_{i\_rtps,n} = \max\{v(A(t)), f_{i\_rtps,n-1}\}$$

$$f_{i\_rtps,n} = S_{i\_rtps,n} + L_{i\_rtps,n} / r_{i\_rtps}$$

$v(A(t))$ is the effectual time of the system; $L_{i\_rtps,n}$ is the length of the arrived packet; r_{i_rtps} is the bandwidth which is assigned for flow number i. In each queue, the packets arrange in non-descending order, and the scheduler will select the packet which has the smallest $f_{i\_rtps,n}$ value.

The scheduler 2 is responsible for scheduling nrtPS with IWFQ algorithm. It neglect the service whose reserved minimum rate is 0.the scheduling principle is similar with schedule 1, and gives more compensation for nrtPS for its delay-insensitive.

Scheduler 3 takes charge of scheduling BE with IWFQ algorithm. BE service flow main bearing messages, e-mail and other business, which has few requirements for delay but more demand for data accuracy. Therefore, under the wicked wireless link circumstances, we should fully guarantee the necessary compensation for the lagged data flow, rather than discard them.

Scheduler 4, as the second-level scheduler, has two functions for different QoS requirements: Firstly, it should assign bandwidth for such fixed cycle service as UGS and ertPS with cycle Scheduling Algorithm; Secondly, it should schedule the queue of scheduler 1, 2 and 3 with WRR algorithm.

For the highest priority of UGS, at the initial allocation of bandwidth, the scheduler firstly assigns bandwidth for the UGS, then turn to ertPS. In their subsequent Scheduling process, BS scheduler will allocate bandwidth by their periods. Under this rule, we get the outcome that, for any service flow i, the time of the n times schedule is:

$$t_{i,n} = t_{i,0} + (n-1)t$$

$t_{i,0}$ presents the first time of authorization; t presents the time interval of periodical schedule.

The second-level scheduler adopt WRR algorithm to deal with the scheduler 1, 2 Scheduler, Scheduler 3. Different weights are set to the first-level scheduler according to the priority of the service flow scheduled by them. Scheduler 1 will get the biggest weight and the scheduler 3 will get the smallest one. Each queue will be set a counter, whose initial value is its weight. The counter will subtract 1 after being scheduled. If a counter's value is 0, the corresponding queue should not to be scheduled. Until all counters decrease to 0, it will reset them to the weight value again. WRR algorithm is able to ensure that the packets with high real-time requirement receive more services.

3.3 Performance analysis

1. Fairness analysis

In terms of IEEE802.16 agreement, UGS, ertPS adopts periodical initiative authorization scheduling method, and the BS dispatch bandwidth for UGS periodically. Clearly, it is a fair scheduling strategy for all connection. For rtPS, nrtPS and BE, in the first-level scheduling process with IWFQ, scheduler chooses the packets which needs smallest division of time each time, and the compensation for delayed flow and punishment for the ahead flow are clearly defined. These options avoid infinite compensation, and help to keep a fair schedule for rtPS, nrtPS and BE.

2. Throughput

Monitoring each connection statue in the link, the scheduler stop to assign bandwidth for connection once it has error. Bandwidth released will be dispatch for other good state connection, which helps to enhance throughput of the system.

3. Delay analysis

It is impossible to exist delay issue as long as UGS and ertPS have a good connection statue, for their bandwidth is assigned compulsively in terms of the needs of connection. By the scheduling algorithm, rtPS selects the packet which has the smallest final time value to service in each schedule period, which makes the service waiting for the longest time can be deal with earlier and ensures the lower delay for the service. nrtPS can tolerant delay. BE, as a kind of best effort service, can ignore transmission delay.

4. Simulation and analysis

In this section, we validate the strategy and the algorithm by building a simulation scenario based on the NS2 platform. NS2, an object-oriented network simulation tool [11], can simulate the entire network environment. It uses a script language-with characteristics of the object-oriented programming language-called OTcl and C + + as a programming language. It runs on the Windows or Unix operation system. All the source codes of NS2 are opened, which is better for network study and expansion. It has so many advantages that lots of researchers fall over their self for NS2. In addition, we choose Gawk Language to analyze the simulation results, as the Gawk programming language has powerful character processing ability.

4.1 Scenario

We set a BS and fifteen SS in the simulation platform, in which every three SS bearing the same class of service, that is, three bear UGS service, three bear ertPS service, three bear rtPS service, three bear nrtPS service and the other three 3 bear BE service. For UGS and ertPS, transmission rate are set 64 kbps; For rtPS, nrtPS and BE, maximum transmission rate are set 1 Mbps.

The purpose of this experiment is to validate whether the admission and the scheduling algorithm can ensure the QoS of the variety services. We will illustrate the paradigm from average delay.

The data results from procedure are stored in "nam" and "trace" documents. The process of information exchanging between BS and SS will be display dynamically, just as fig 4.1.

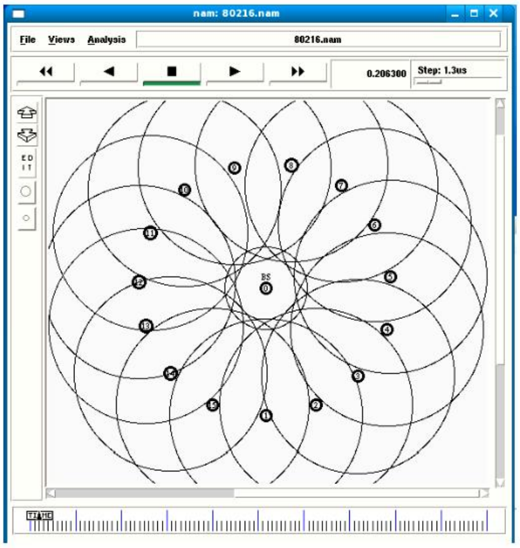

Fig 4.1 Process of information exchanging between SS and BS

4.2 Analysis of trace file

A mass of data results from procedure are stored in "trace" documents, which reflects the entire process form SS registration to data transmission. Parts of them present as follow:

s 0.003861330 _5_ MAC --- 0 RNGREQ 20 [0 5000000 0 0] ------- [0:0 0:0 0 0]
s 0.003861330 _1_ MAC --- 0 RNGREQ 20 [0 1000000 0 0] ------- [0:0 0:0 0 0]
s 0.003861330 _3_ MAC --- 0 RNGREQ 20 [0 3000000 0 0] ------- [0:0 0:0 0 0]
s 0.003861330 _2_ MAC --- 0 RNGREQ 20 [0 2000000 0 0] ------- [0:0 0:0 0 0]
s 0.003861330 _4_ MAC --- 0 RNGREQ 20 [0 4000000 0 0] ------- [0:0 0:0 0 0]

The first row of the data presents the behavior of the event, in which s presents packets be sent; r presents packet be received; D presents packets be discarded. The second row presents the time of the event. The third row presents the number of the nodes. The fourth row presents whether it is MAC packet or RTR packet. The sixth row presents the number of the packet. The seventh row presents the class of the packet. The last row presents the size of the packet. For example, the first line of the data explains that SS set a registration message to BS at 0.003861330s.

After analysis by Gawk language, the result can be displayed in Gnuplot-a drawing tool-as follow.

It displays the average delay Stat. results of the services which has high real-time requirement in fig 4.2. In terms of the results, we can see that average delay of the UGS is about 4ms, 42ms for ertPS and 57ms for 57ms. The delay for these three classes of services can ensure their QoS.

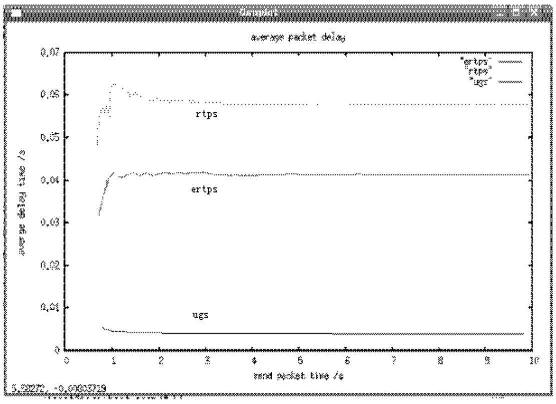

Fig 4.2 Average delay of UGS、ertPS、nrtPS

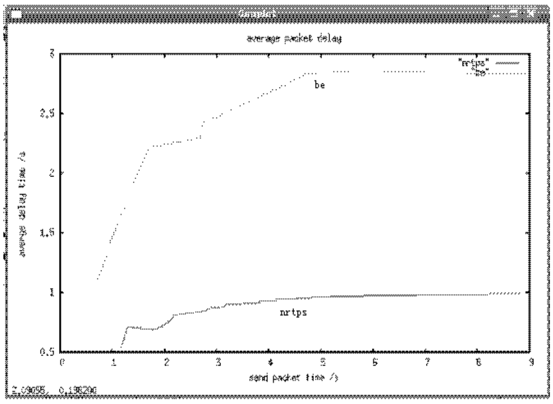

Fig 4.3 Average delay of nrtPS、BE

Fig 4.3 shows the average delay Stat. results of nrtPS and BE services which has average real-time requirement. From the result, we can see that the average delay of nrtPS is 1s, and 3s for BE, which can satisfy their QoS requirements.

Integrating fig 4.2 and fig 4.3, we can conclude that both service with high priority and one with low priority can be satisfied by the dmission control strategy and the scheduling algorithm we proposed in this paper.

5. Conclusion

There is no specific definition of admission control and service scheduling algorithm for MAC layer QoS in IEEE802.16 standard, therefore, those aspects of QoS guarantee mechanism attract more attention of the researchers. In this paper, according to the characteristics of the different services, we firstly propose a new admission control strategy, which is based on the minimum reserved bandwidth. Meanwhile, we describe the admission process respectively. Secondly, we show a multi-level scheduling algorithm to support the admission control strategy. Finally, we validate the effectiveness of the strategy and the algorithm by a simulation experiment.

Our study mainly based on the PMP mode of communication structure between BS and SS, however, the Mesh mode of communication structure has not been studied in this paper. As well as, we did not consider the switching issue result from mobile devices span different two areas. These will be our future work.

References

[1] Xun Yang, Gang Feng, David Siew Chee Kheong, "Call Admission Control for Multservice Wireless Networks with Bandwidth Asymmetry between Uplink and Downlink," *IEEE Transactions on Vehicular Technology*, Jan.2006, vol.55, no.1.

[2] Xiaojun Ma，Guanqun Gu，An Adaptive Measurement-Based Admission Control Algorithm，*Chinese Journal of Computers*，2001，Vol 24，no1，pp.40-45.

[3] Chang Wook Ahn, R.S.Ramakrishna, "QoS Provisioning Dynamic Connection-Admission Control for Multimedia Wireless Networks Using a Hopfield Neural Network," *IEEE Transactions on Vehicular Technology*, 2004, vol.53, no.1, pp.106-117.

[4] L.Badia, M.Zorzi, A.Gazzini, "A Model for Threshold Comparison Call Admission Control in Third Generation Cellular Systems," *Proc.IEEE ICC'03*, May.2003, vol.3, pp.64-68.

[5] Jinwen Zhang, etc.，802.16 Wireless MAN Broadband Access Technology，*Publish house of electronics industry*，April, 2006.

[6]H.Fattah, C.Leung, "An Overview of Scheduling Algorithms in Wireless Multimedia Networks," *IEEE Wireless Communications*, Oct.2002, vol.9, no.5, pp.76-83.

[7] J.R.Bennett, H.Zhang, "WF2Q:Worst-case fair weighted fair queueing," *IEEE INFOCOM'96*, Mar.1996, vol.1, pp.12-18.

[8] Yaxin Cao, Victor O.K.Li, "Scheduling Algorithms in Broad-Band Wireless Networks," *Proceedings of the IEEE*, Jan.2002, vol.89, no.1, pp.76-87.

[9] K.Wongthavarawat, A.Ganz, "Packet Scheduling for QoS Support in IEEE802.16 Broadband Wireless Access System," *International Journal of Communication Systems*, 2003, vol.16, pp.81-96.

[10] S.Lu, V.Bharghavan, R.Sirkant, "Fair Scheduling in Wireless Packet Networks," *IEEE/ACM transactions on networking*, Aug.1999, vol.7, no.4, pp.473-489.

[11] Ji qin，Xuesong Jiang，Mobile IP Technology and NS2 Simulation，*China machine press*, October 2006, pp.85-21.