國立交通大學

資訊科學與工程研究所

碩士論文

IEEE 802.16e 無線寬頻存取系統下的電源節省管理及 QoS 排程機制

A Joint Design of Power Saving and QoS Scheduling in IEEE 802.16e Broadband Wireless Access Systems

研究生：林佳燕

指導教授：曾文貴 教授

聯合指導教授：趙禧綠 助理教授

中華民國九十六年九月

IEEE 802.16e 無線寬頻存取系統下的電源節省管理及 QoS 排程機制

A Joint Design of Power Saving QoS Scheduling in IEEE 802.16e Broadband Wireless Access Systems

研 究 生 ： 林佳燕　　Student: Chia-Yen Lin

指 導 教 授 ： 曾文貴　　Advisor: Wen-Guey Tzeng

聯合指導教授： 趙禧綠　　Co-Advisor: Hsi-Lu Chao

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

碩 士 論 文

A Thesis
Submitted to Institute of Computer Science and Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
In

Computer Science and Engineering

Sep 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年九月

摘要

IEEE802.16e 是最近熱門探討的技術，其標準定義了媒介存取控制層和實體層的協定，其中實體層的頻段是介於 10~66GHz 以及低於 11GHz，而媒介存取控制層支援分時雙工和分頻雙工這兩種模式，並且定義了五種服務類別來支援服務品質，這五種等級分別是：UGS, rtPS, nrtPS, BE, and ertPS，但是 802.16e 並沒有定義針對不同服務類別的排程機制，此外 IEEE 802.16e 亦支援移動性，所以如何有效利用電源成為一重要研究議題。

本篇論文是在 IEEE 802.16e 上提出電源節省機制的架構，包括了電源控制的策略以及排程方法，整個機制是建構在分時雙工的模式上，在電源控制策略上是希望利用集中傳送時間的方式，以減少需要開啟電源的時間，達到省電的效果又能夠滿足其服務品質，並且提出一個公平的排程機制來合理的分配頻寬。

本篇論文透過模擬的方式來評估所提出演算法的效能。模擬的結果顯示提出之機制確實能夠達到省電的效果且公平分配頻寬。

Abstract

The IEEE 802.16e standard for broadband wireless access is a recently popular technology which defines the medium access control (MAC) layer and the physical (PHY) layer with frequency bands between 10 and 66GHz and below 11 GHz. The MAC layer supports two duplexing, i.e., time division duplexing (TDD) and frequency division duplexing (FDD), and supports QoS by defining five service classes. However, how to do scheduling among these service classes is not defined in the standard. In addition, the IEEE 802.16e supports mobility, and thus how to utilize battery efficiently is an important issue for protocol design.

In this thesis, we propose a novel power saving framework which includes a transmission merging mechanism and a scheduler for the IEEE 802.16e. This framework is designed for TDD mode. The transmission merging mechanism focuses on how to merge the transmission time to reduce the number of awake-frames. To still satisfy each connection's quality of service (QoS) demand, and achieve fairness, we design a scheduler to allocate downlink (DL) and uplink (UL) subframes.

In this thesis, we evaluate the proposed algorithm by simulations. The results show that the proposed method achieves both power efficiency and fair BW allocation.

致謝

經過這兩年的努力，終於把論文順利的完成了，這一路走來好像完成了一個不可能的任務似的，以前看學長們辛苦後開心的樣子，到現在自己終於能體會那種解脫和滿足的感覺，雖然還是覺得這一切都有些不可思議，但是我竟然還是做到了。

這兩年來要感謝的人很多，最感謝的是我的指導老師—趙禧綠教授，讓我這個有些小懶散的學生，在她的鞭策、鼓勵之下，仍舊每天都能夠有一點一滴的成長，也才能完成這篇論文，不管是生活上還是學業上，因為有她的關心和包容，才能讓我順順利利的從研究所完成學業，真的非常謝謝她。

還要感謝的是實驗室夥伴們，不管是課業或研究上有什麼問題，他們都會跟我一起討論、一起解決，在生活上，大家互相安慰、互相鼓勵，有他們的陪伴，讓我在實驗室過了非常開心的兩年；此外，也要感謝大學時代的同學們，雖然大家分屬在不同的實驗室，但是一旦碰到什麼問題，他們都還是會主動的關心我、鼓勵我，讓我深深感受交大是一個很有人情味的地方。

最後我要感謝我的家人，雖然她們沒辦法在課業上幫我些什麼，但是因為有她們鼓勵和期望，讓我有動力繼續努力；有這麼多人的陪伴，我真的很幸福，謝謝你們！

Contents

List of Figures

List of Tables

Chaper 1 Introduction

The IEEE 802.16 broadband wireless access standard called Worldwide Interoperability for Microwave Access (WiMax) which is developed by the IEEE 802.16 working group on broadband wireless access was proposed for wireless metropolitan area networks (MANs). It is important and popular because of providing data transmission over long distances and high data rates over a large area to a large number of users. This could result in low pricing for both home and business customers. This standard specifies the air interface of a fixed point-to-multipoint (PMP) broadband wireless access system which provides broadband wireless access between subscriber stations (SS) such as residential or business customers and a base station (BS) such as the internet service provider (ISP). The 802.16d [1] was proposed in 2004 which is a called 802.16-2004, too. The IEEE 802.16e [2] was proposed in 2006 which support mobility and power saving. In IEEE 802.16e, the mobile subscribe stations are abbreviated as MSSs.

[1, 2] specifies MAC and PHY layers. The PHY specifies the frequency band between 10 and 66 GHz and below 11 GHz. The frequency band between 10 and 66 GHz requires line-of-sight (LOS), and the one below 11 GHz requires near-LOS and non-LOS (NLOS). Two duplexing, i.e., TDD and FDD are supported by the defined MAC protocol.

1.1 Service Classes in IEEE 802.16e

The MAC supports QoS by defining five service classes, i.e., Unsolicited Grant Service (UGS), Real-time Polling Service (rtPS), Non-real-time Polling Service (nrtPS), Best Effort (BE) and, Extend Real-time Poling Service (ertPS). Each is described in the following.

(1) UGS: it is designed to support real-time data streams consisting of fixed-size data packets, and packets are periodically. Examples include T1/E1 and Voice over IP (VoIP) without silence suppression.

(2) rtPS: it is designed to support real-time data streams consisting of variable-sized data packets, and packets are periodically. An example includes moving pictures experts group (MPEG) video.

(3) nrtPS: it is designed to supports delay-tolerant data streams consisting of variable-sized data packets and a minimum data rate is required. An example includes FTP.

(4) BE: it is designed to support data streams, and no minimum service level is required. It therefore may be handled on a space-available basis.

(5) ertPS: it is newly added in the 802.16e. It is a scheduling mechanism which builds on the efficiency of both UGS and rtPS. The BS shall provide unicast grants like UGS manner and dynamic allocation like rtPS.

In this thesis, we do not consider ertPS connections.

1.2 An Overview of Power Saving Mechanisms

Power saving is a very important issue in wireless network. It can be solved by physical component improvement or MAC protocol redesigns. In this thesis, we focus on MAC protocol redesigns. Existing algorithm can be classified into two common solutions. The first kind of solutions is adjusting transmission ranges to reduce consuming power and interferences. The second kind of solutions is allowing a host to turn off its antenna when no packets waiting to be transmitted.

For the first kind of solutions, many exiting solutions are designed for multihop related network. In recent years, many people have an upsurge of interest in multihop related network. By adding related nodes, the station can use less power and raise data rate. In [3], the authors proposed two dynamic coverage control schemes of relay systems for effective power saving scheme, which can provide flexible cell coverage variation with minimal usage of power consumption.

For the second kind of solutions, exiting solutions try staggering transmissions to avoid collisions or to get chances to turn off antenna to save power. In [4], the authors proposed a power controlled multiple access wireless MAC protocol within the collision avoidance framework, called PCMA. PCMA improves the throughput performance. In [5], the author proposed a fair queuing algorithm with power saving. The algorithm is to find out the packet deliver sequences with minimum power consumptions. In [6], the author proposed a framework to implement power managers, definition of user perceived performance, and quantitative comparison of algorithms. It shows the minimum length of an idle period to save energy by some effects, such as energy to shut down, and to wake up, and so on. For 802.16e, it saves power by using this kind of solution. It proposes a new power saving scheme. It shows a new item called power saving class which classifies different connections. We will introduce in next section.

1.2.1 Power Saving Classes in IEEE 802.16e

In IEEE 802.16e systems, a MSS can enter the sleep mode to reduce its energy consumption. Sleep mode is intended to minimize power usage and decrease usage of serving BS air interface resources. For each involved MSS, the connections will be classified into power saving classes. A power saving class is a group of connections with common demand properties. Each power saving class may be repeatedly activated and de-activated. When a power saving class is activated, it means this power saving class starts sleep/listening windows sequence. For a MSS, unavailability intervals are the time intervals that don’t overlap with listening windows of any active power saving classes, and availability intervals are the time intervals that don’t overlap with unavailability intervals. During unavailability intervals, the BS shall not transmit to the MSS, so the MSS may turn off some physical operation components to reduce power consumption. During availability intervals, the MSS receives all DL transmissions same way as in the state of normal operations.

There are three types of power saving classes defined in IEEE 802.16e. Type I is recommended for connections of BE and nrtPS. Type II is recommended for UGS and rtPS. Type III is recommended for multicast connections as well as for management operations.

As shown in Figure 1.1, there are two power saving classes. The union of listening windows of power saving classes is the availability intervals, and the intersection of sleep windows of power saving classes is the unavailability intervals. The MSS can save power only in unavailability intervals.

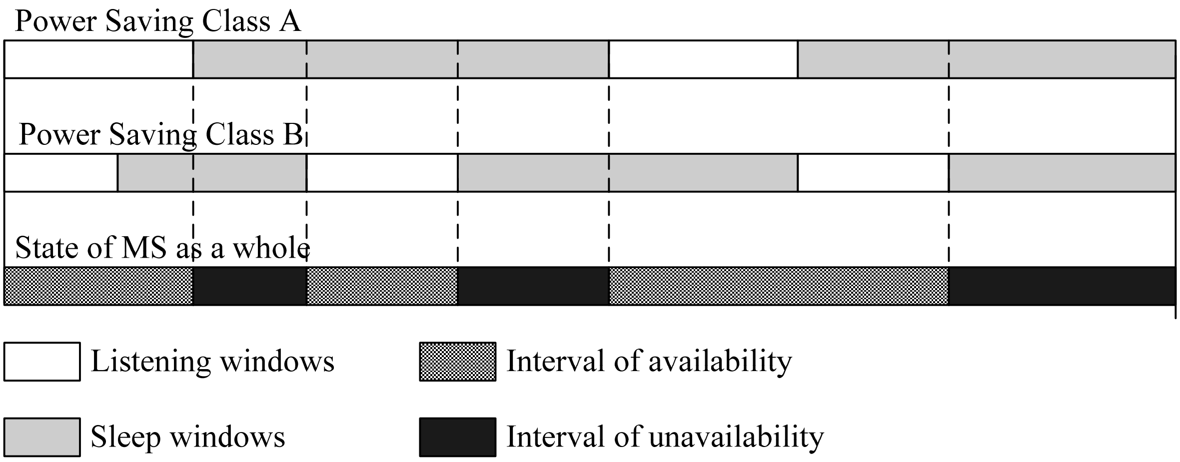

Figure 1.1 An example of sleep mode operation with two power saving classes

1.3 Motivation

In 802.16 standard, the MAC supports QoS but there are no a packet scheduling solution. Because the power saving scheme in IEEE802.16e focuses on one power saving class, it is less effective in saved power of one node. In this thesis, we will propose a novel power saving framework which includes a transmission merging mechanism and a scheduler for the IEEE 802.16e. This framework is designed for TDD mode. The transmission merging mechanism focuses on how to merge the transmission time to reduce the number of awake-frames. To still satisfy each connection's QoS demand, and achieve fairness, we design a scheduler to allocate DL and UL subframe.

The remainder of this thesis is organized as follows. In chapter 2, we introduce related work. In chapter 3, we describe our proposed scheduling algorithm with QoS support and power saving. In chapter 4, we show the performance evaluation. In chapter 5, we conclude this thesis and give the future work

Chaper 2 Related Work

In [7], the authors proposed fair and efficient service flow management architecture for IEEE802.16. It applies Deficit Fair Priority Queue (DFPQ) for different service classes. Using the DFPQ scheduling overcomes the higher priority connections starve the bandwidth (BW) of lower priority connections. The DFPQ uses two parameters, i.e., “quantum” and “DeficitCounter” to adjust satisfaction degree of a connection and to avoid higher priority connections gets too much BW affecting the low priority connections’ transmission opportunity. In this paper, it only proposed a scheduling algorithm but didn’t consider about power saving.

For power saving in 802.16e, [8][9][10][11] analyzes the binary-increasing sleep window of the type I power saving class by proposing a mathematical model, but doesn’t consider the type II power saving class. In [12], the authors provided a mathematical analytical model which is capable of calculating the power efficiency and packet access delay for Type I and Type II power saving classes. For type I power saving classes, they uses the embedded Markov chain model to represent the binary-increasing window of sleep state and for type II power saving classes, they uses two-state Markov chain model to represent the constant sleep window size. They show that Type I power saving classes can get good power saving performance, but have a dramatic negative effect on packet delay because of binary-increasing sleep behavior. Type II power saving classes can maintain good packet delay, but they consume more power because of periodic awaking behavior.

In [13], the author proposes two scheduling algorithms for sleep mode operations to move the transmissions of Type II power saving classes. It maximizes the sleep period of

Type II power-saving without violating QoS of the connections, For the first scheduling algorithm, it proposed a periodic on-off scheme for one connection which distributes small packets to all OFDM frames , and groups small packets in the less number of OFDM frames under all connections' QoS requirements achievement. However, as shown in the Figure, 2.1, if the Delay is smaller then the grant duration, it can be merged.

For the second scheduling algorithm, it proposed an aperiodic on-off scheme (AS) which schedules the packets in the minimal number of frames and also guarantees the QoSs for many connections. It merges the transmission into fewer frames without violating the delay constraint of different connections. However, the algorithm just finds fewer awake-frames, but doesn't consider how to find longer sleep interval. The longer sleep interval can reduce the power consumption of components' on/off. As shown in Figure 2.2, the algorithm will merge the transmissions of the first connection and second connection into the i^{th} frame. But, if merging them into j^{th} frame, it can get longer sleep interval with less components' on/off.

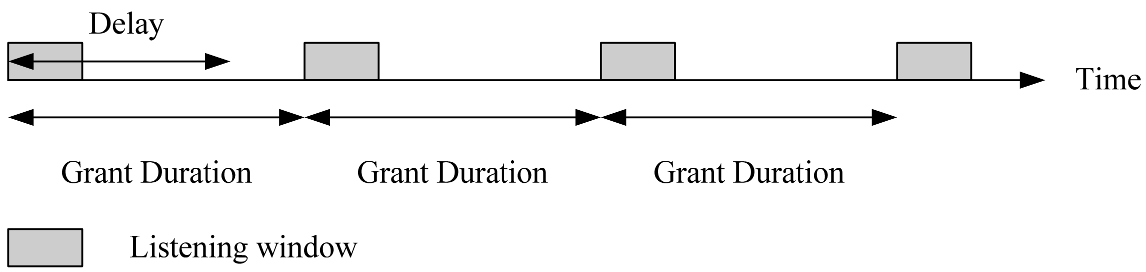

Figure 2.1 A problem of periodic on-off scheme

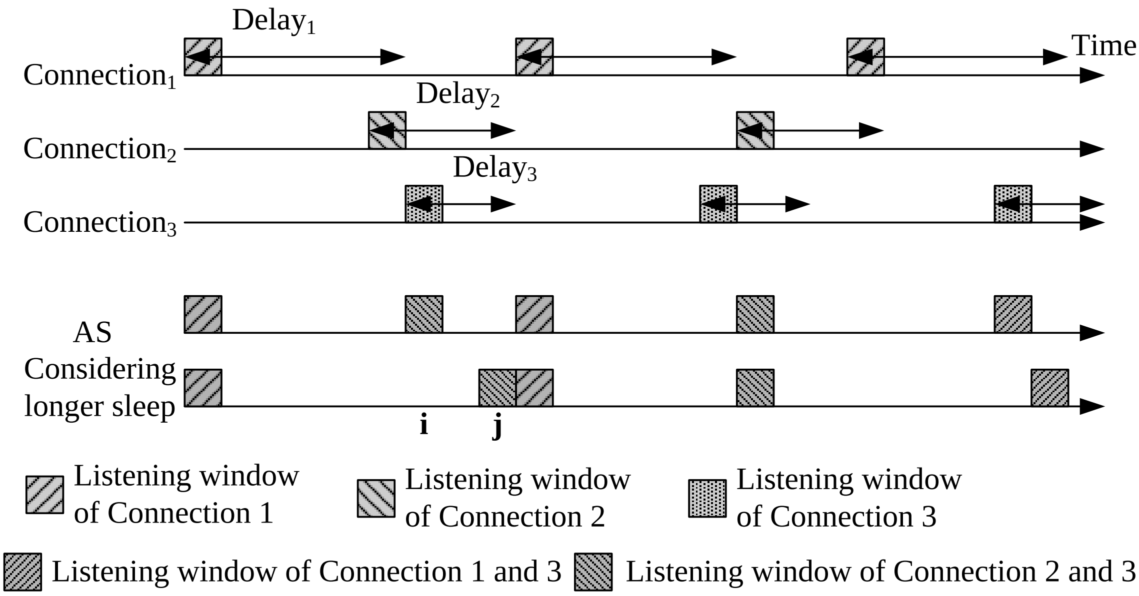

Figure 2.2 A problem occurs when $Delay_3 < Delay_2 < Delay_1$ in AS

Moreover, for these two schemes, they only consider the periodic fixed size traffic which can be known before, but the variable packets size, like rtPS is not considered. The rtPS connections' queue size information is obtained by polling or piggyback, and is know in the run- time. Otherwise, it also doesn't consider the QoS requirements of Type I power saving class.

Chaper 3 The Proposed Scheduling Algorithm with QoS Support and Power Saving

In this chapter, we propose a novel power saving framework which includes a transmission merging mechanism and a scheduler for the IEEE 802.16e.

3.1 System Environment

We construct our system in TDD mode, and just consider four service classes, i.e., UGS, rtPS, nrtPS, and BS. For UGS connections, the parameter p_i, v_i, and d_i should be specified. The d_i is maximum latency. The p_i is grant duration. The v_i is the fixed amount of BW (BW). It means that the BS should allocate v_i bytes in every p_i frames. The maximum grant latency is d_i frames. For rtPS connections, the parameter r_{min}, r_{max}, and d_i should be specified. For nrtPS connections, the parameter r_{min}, and r_{max} should be specified. The r_{min} means minimum reserved rate. The r_{max} means maximum sustained rate. The d_i means maximum latency. The detail is in the Table 3.1. Here the unit of l_i and d_i is “frames”. It’s because the frame duration is very small and sometimes the actually transmission point is not easy to predict. If the Jitter is 51msec and the Frame duration is 5msec, the Jitter will be set $\lceil 51/5 \rceil$ =10 frames.

Table 3.1 The parameters of service flow

	Name	Service Flow	Unit
p_i	grant bandwidth	UGS	Bytes
l_i	grant duration	UGS	Frames
r_{min}	minimum reserved rate	rtPS, nrtPS	Bytes/sec
r_{max}	maximum sustained rate	rtPS, nrtPS, BE	Bytes/sec
d_i	maximum latency	UGS, rtPS	Frames

The admission control determines whether a new flow can enter the system by using the DSA, DSC, and DSD messages. The BS uses the admission control policy to decide whether the QoS request of a connection can be satisfied. In this thesis, (3.1) is the admission control policy. The policy is very simple. The idea is the total minimum requirement must be smaller then 80% Capacity.

$$\sum_{i-0}^{I}\frac{v(i)}{l(i)\times D}+\sum_{j=0}^{J}r_{min}(j)+\sum_{k=0}^{K}r_{min}(k)\le C\times 0.8, \qquad (3.1)$$

where C means the channel capacity, and D means the frame duration. I means the total number of UGS connections, J means the total number of rtPS connections, and K means the total number of nrtPS connections. The UGS connection i uses v(i) as the grant BW, l(i) as grant duration. $r_{min}(j)$ means r_{min} of rtPS connection j. $r_{min}(k)$ means r_{min} of nrtPS connection j. $\sum_{i-0}^{I}\frac{v(i)}{l(i)\times D}$ means the summation of the UGS connections' rate. $\sum_{j=0}^{J}r_{min}(j)$ means the summation of the rtPS connections' r_{min}. $\sum_{k=0}^{K}r_{min}(k)$ means the summation of the nrtPS connections' r_{min} .

3.2 System Model

All MSSs in the system are classified into power-saving mode (PS mode) and non-power-saving mode (non-PS mode). Here, a MSS in PS mode is called a PS MSS, and a MSS in non-PS mode is called a non-PS MSS. Initially all MSSs are in non-PS mode. A MSS can adjust the self-condition to decide whether sending entry PS mode request to its BS or not. Once a MSS is permitted to enter PS mode, the transmission of the MSS will be merged in order to reduce the consumed power. The BS re-schedules and centralizes transmission time of the MSS in certain frames according to the QoS parameters of the MSS's connections. Then the BS informs the MSS of the next awake-frame by sending map. When the MSS receives the map, it will know how many incoming frames it can go sleep to save power. If a MSS wants to exit PS mode, it can send exit PS mode request to BS. If the BS is busy, it can take the initiative in asking the MSS exiting PS mode.

The Figure 3.1 shows the system model. The phase one is for a MSS to decide weather sending entry/exit PS mode request. The phase two is for BS to permit entry PS mode requests from MSSs and accept exit PS mode requests. In addition, BS can take the initiative in asking the MSS exiting PS mode in this phase. The phase three introduces how to generate a frame, merge PS MSSs' transmission, and schedule the services. We will describe the detail of each phase in following subsection.

When a MSS enters PS mode, the transmission of the MSS will be merged into fewer awake-frames, as shown in Figure 3.2. The sleep intervals mean the duration between two awake-frames. Our aim is to find less awake-frames under QoS satisfaction and get longer sleep intervals.

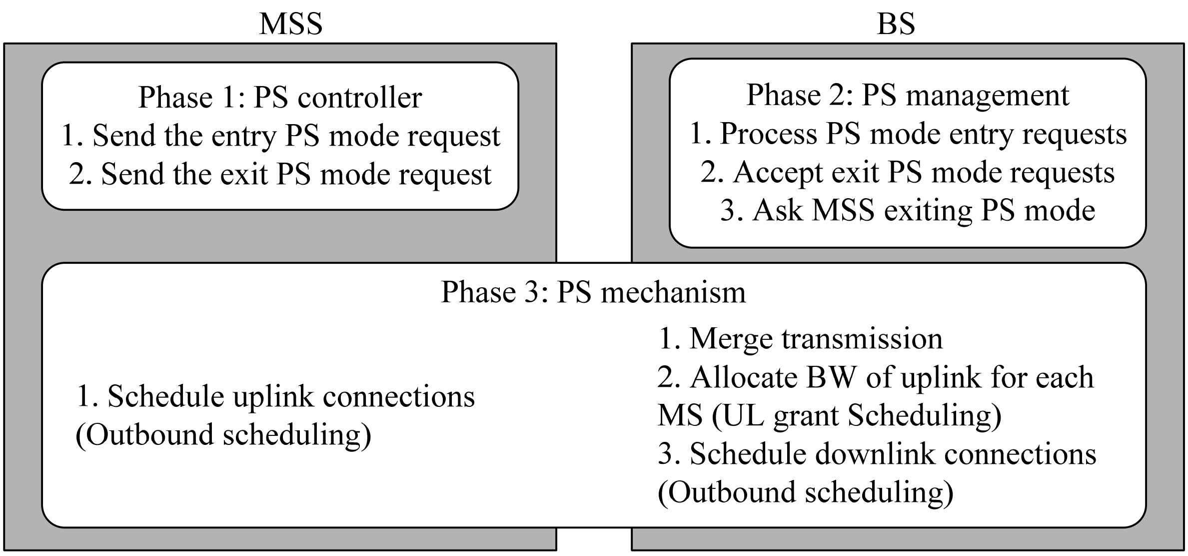

Figure 3.1 System Structure

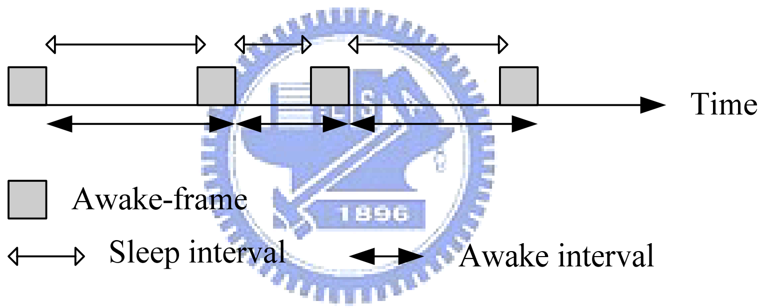

Figure 3.2 Transmission behavior of PS mode

3.2.1 Phase 1: PS Controller

The Phase 1 is executed at MSSs as shown in Figure 3.3. A MSS would periodically monitor its condition. If the traffic load is light, it will request to enter PS mode by sending MOB_SLP-REQ message. The BS replies the request by sending MOB_SLP-RSP message. Weather sending request , there are some factors：

(1) The number of connection;

(2) The connection’s traffic load; and

(3) Uplink queue size.

When a MSS is in PS mode and the traffic load is heavy, it can request to exit the PS mode by sending MOB_SLP-REQ message. The BS will always accept this request, and sends MOB_SLP-RSP message. Otherwise, if a MSS gets MOB_SLP-REQ message to be asked to exit PS mode from the BS, it will always accept this request and sends MOB_SLP-RSP message. When there are the following condition, a MSS can consider weather sending a exit PS mode request：

(1) The new flow is added; and

(2) The MSS wants doing handoff.

3.2.2 Phase 2: PS Management

The Phase 2 is executed at BS as shown in Figure 3.3. When BS gets a MOB_SLP-REQ message which is requesting to enter PS mode, it will measure the MSS condition and self-condition to determine whether permitting the request or not by sending a MOB_SLP-RSP message. If BS accepts the request, it will send response with starting point which indicates the PS mode starting frame. Weather permitting the request, there are some factors：

(1) The number of connection;

(2) Downlink queue size; and

(3) BS is busy or not (e.g., the number of PS MSS).

If the message is requesting to exit PS mode, a BS will always accept the request and send the MOB_SLP-RSP message. If the BS's traffic load is heavy, it can ask the MSS to exit PS mode by sending MOB_SLP-REQ message. A MSS always accepts this request and sends MOB_SLP-RSP message. When there are the following condition, the BS can consider weather sending exit PS mode request：

(1) The new MSS is added; and

(2) The new flow is added.

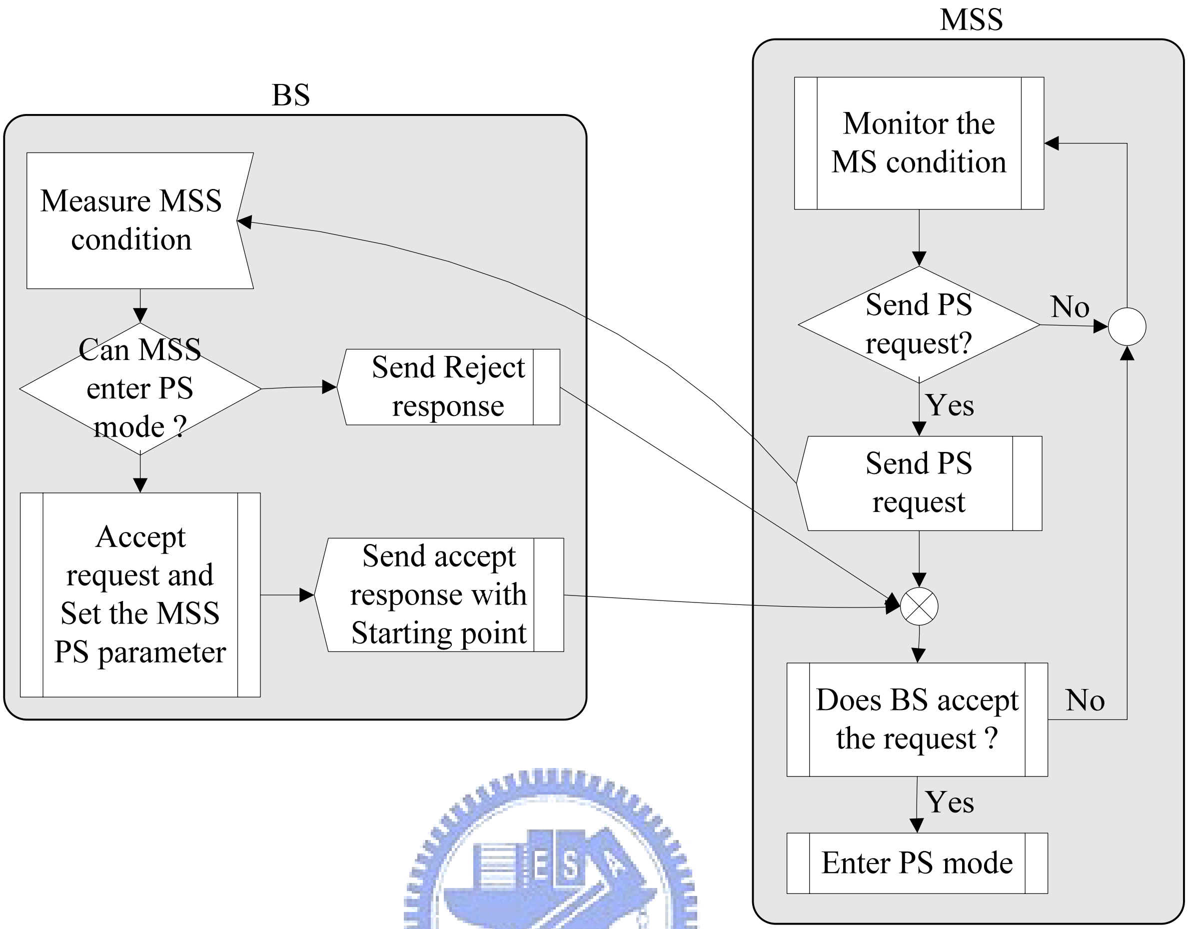

Figure 3.3 Entrance decision of PS mode

3.2.3 Phase 3: PS Mechanism

As shown in the Figure 3.4, when generating a map, the BS calls the transmission merging machine to merge transmission of the PS MSS and calculate the next awake-frame. After calculating the next awake-frame for PS node, the BS calls the Uplink Grant Scheduler to allocate BW for Uplink. Here the allocation is given per MSS. After deciding the allocation of uplink per MSS, the BS will call the Outbound Scheduler to decide the allocated BW per downlink connection. Then, DL/UL map will be generated. The BS will notify each PS MSS the next awake–frame in map. When a MSS receives the map and gets the Uplink BW, it will call the Outbound Scheduler to allocate BW to each uplink connection. In this thesis, we propose a transmission merging mechanism to do merging process, a scheduling strategy to allocate BW.

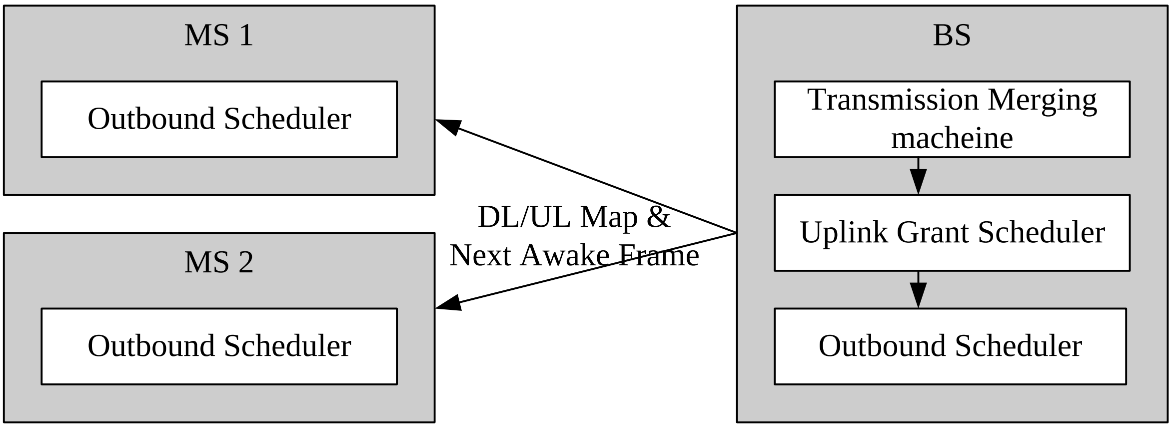

Figure 3.4 Map generation flow

3.3 The Transmission Merging Mechanism and Scheduling Strategy

The transmission merging mechanism is to do merging process, and the purpose is to merge the transmission time into fewer awake-frames and get longer sleep intervals to save power for each PS MSS. The scheduling strategy purpose is to allocate BW fairly

.

3.3.1 Transmission Merging Mechanism

In order to merge the transmission, the first problem is UGS connections which are periodically allocated fixed amount of BW. For UGS connections, the strategy is that using the parameter" maximum latency" adjusts the transmission time. After merging UGS, the latency of the rtPS will be considered to adjust the transmissions time. Finally, the non-real-time services are transmitted along with the real-time services. The major problem of the merging is that how many frames' traffic can be merged into one frame and doesn't exceed the frame size. The detail is below.

(1) UGS connections merging

A UGS connection is periodically allocated BW and each UGS would be given a QoS parameter "maximum latency". If the maximum latency of the connection i of node k is given $d_{i,k}$ frames, it means the allocation can be delayed $d_{i,k}$ frames. So according this parameter, we can merge the UGS allocation of node k to reduce the awake-amount of frames.

The algorithm is as Figure 3.5. By this UGS merging algorithm, we can calculate the farthest next awake-frame frame *F* and the nearest next awake-frame frame *Y*. In our algorithm, we expect to merge transmissions as much as possible. By *Y* and *F*, we can get the adjustment space after UGS merging.

Firstly, it sorts the UGS connections' transmission frames according $Y_{i,k}$ and $F_{i,k}$ which are the next transmission frame, and the farthest transmission frame of the UGS connection i. A UGS connection's transmission can be merged if it satisfies the condition $F_a \leqq F$, where the F_a is the farthest next awake-frame frame of this connection. When a UGS connection is added into the next UGS transmission frame, the *F* and *Y* will be adjusted. The algorithm simple example is shown in the Figure 3.6. The firstly coming transmission is UGS_1, and the transmission of both UGS_2 and UGS_3 can be merged with UGS_1. But, after the transmission of UGS_2 is merged with UGS_1, the transmission of UGS_3 will not be merged with transmission of UGS_1 and UGS_2.

Function: Get_Next_UGS
Input:
For each UGS connection **i** of PS mode MSS **k**

$p_{i,k}$: UGS Duration(frames), $d_{i,k}$: Max latency(frames), $v_{i,k}$: BW-allocated

$Y_{i,k}$: The next transmission frame of the UGS i

$F_{i,k} = Y_{i,k} + d_{i,k}$

There are N UGS connections of i.

Output:
Y : Farthest next awake-frame F : Nearest next awake-frame

1. Sort the N connections according the $Y_{i,k}$,and $F_{i,k}$, denoted
   $S = \{(Y_1, F_1), \ldots\ldots, (Y_N, F_N) \mid Y_i < Y_j \text{ or if } (Y_i = Y_j)\ F_i < F_j \text{ for any } i<j\}$
2. int a=2; Y= Y_1; F= F_1;
   while(a≦N){
     if($F_a \leq F$){ //merge
       F = Min (F , F_a); //the farthest merge point
       Y = Max (Y , Y_a); //the nearest far merge point
       a++;
     }
     else
       break;
   }

Figure 3.5 The algorithm of UGS merging

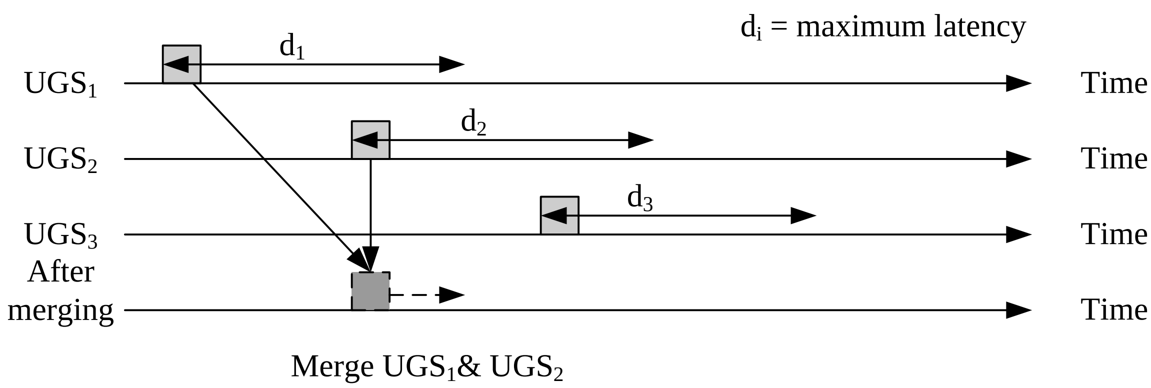

Figure 3.6 The algorithm illustration of UGS merging

The Figure 3.7 shows an example with three UGS connections. The parameters of each connection is shown in Table 3.2. In the example, after UGS merging, we can see the 5th frame of UGS 3, the 6th frame of UGS 2, and the 7th frame of UGS 1 can be merged into 7th frame. The 9th frame of UGS 2 and the 10th frame of UGS 3 can be merged into 10th frame. The 14th frame of UGS 1, the 15th frame of UGS 2, and the 15th frame of UGS 2 can be merged into 15th frame. There is some adjusting space for 3rd frame and 12th frame. The adjusting space is considered in rtPS merging

Table 3.2 Merging example parameter table

	Grant Duration(frame)	Max Latency(frame)
UGS 1	5	1
UGS 2	3	1
UGS 3	9	2

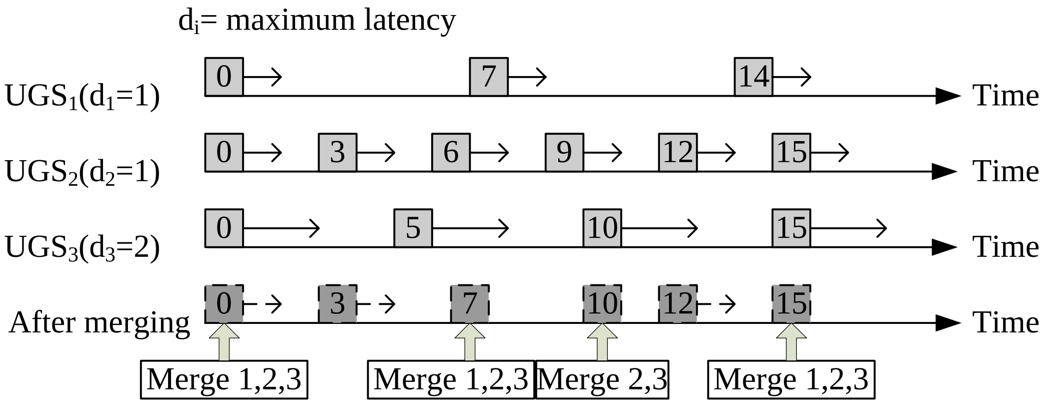

Figure 3.7 An example of UGS transmission merging

(2) rtPS connections merging

An uplink rtPS connection is periodically polled by the BS or updates packet information by piggyback. In this system, the BS gets the rtPS queue size by piggyback for PS mode MSS. If the awake interval is too long, there will be two problems. One is the maximum latency parameter is not be satisfied and the other one is the merging of the rtPS connections maybe exceed the frame size.

For the first problem, as shown in the Figure 3.8, the packets coming between i^{th} and $(i+1)^{th}$ awake-frame are known by BS when node uplink data in the $(i+1)^{th}$ awake-frame by piggyback. These arriving packets must be scheduled in d which is the maximum latency. So, the $(i+2)^{th}$ awake-frame must be appear in d. In order to satisfying all rtPS of the MSS, we choose the minimum value. So, for the first problem, we can get two formulas：

1. The maximum awake interval is bounded by $n_1(k) = \text{Min}\ (d_{i,k})-2$ for MSS k; and
2. Any three awake-frames must be less than Min $(d_{i,k})$,

where $d_{i,k}$ is the delay of the i^{th} connection for the MSS k

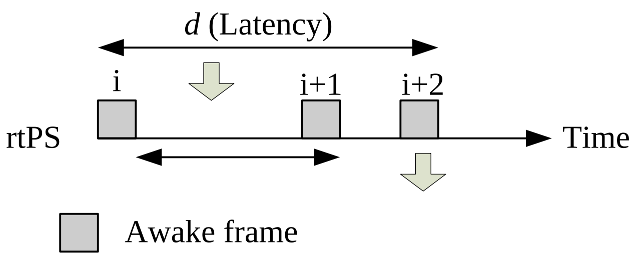

Figure 3.8 The relation between awake-frames and maximum latency of a rtPS connection

For the second problem, we can consider this problem by rtPS and nrtPS's r_{min}. We want to find how many frames' traffic of the MSS can be merged into one frame. Firstly, we calculate the residual capacity by the essential requests of all connections, as shown (3.2).

$$\text{Residual\_C} = C - \left(\sum_{\text{All UGS}} \frac{v(i)}{l(i) \times D} + \sum_{\text{All rtPS}} r_{min}(i) + \sum_{\text{All nrtPS}} r_{min}(i) \right), \tag{3.2}$$

Then, we calculate the average BW request of the PS MSS k.

$$\text{Avg}(k) = \sum_{\text{rtPS of MSS k}} r_{avg,k} + \sum_{\text{nrtPS of MSS k}} r_{avg,k} \tag{3.3}$$

We can get the maximum awake interval as below.

$$n_2(k) = \frac{\text{Residual\_C}}{\text{Avg}(k)} + 1 \tag{3.4}$$

The item $\frac{\text{Residual\_C}}{\text{Avg}(k)}$ means how many other frames' traffic of the PS MSS can be merged into this frame.

In order solving these two problems, we get the maximum awake interval as

$$n(j) = \text{Min}(n_1(j), n_2(j)) \text{ for MS j} \tag{3.5}$$

The merging algorithm for UGS and rtPS is shown in Figure 3.9. The main idea of the algorithm is to consider the rtPS latency requirement, and the maximum awake interval and to decide the next awake frame with longest sleep interval according to the *Y* and *F*. If necessary, it will add extra awake-frame to make it.

Function：Get_Next_Frame

Input：

D_k= Min($D_{i,k}$) –minimum value of all rtPS connections' latency of the MSS k.

$D_{i,k}$ is the latency of i^{th} connection of MSS k

n - The max awake interval,

Output：

Y - next_awake_frame #

1. Call Get_Next_UGS() to get the *Y* and *F*
2. if(($Y \leqq n$) && ($Y \leqq D_k$))

   Compare the *Y* and *F* with next UGS frame, and choose better one

   Output *Y* or *F*

   else

   Output F = Min(D_k,-1, *n*);

Figure 3.9 The algorithm of UGS merging considering rtPS connections

Figure 3.10 is the example after adding the rtPS connection. The minimum latency of all rtPS connection "d" is 7 frames and the maximum awake interval "n" is 3. So, it needs to add awake-frame at the 6^{th} frame

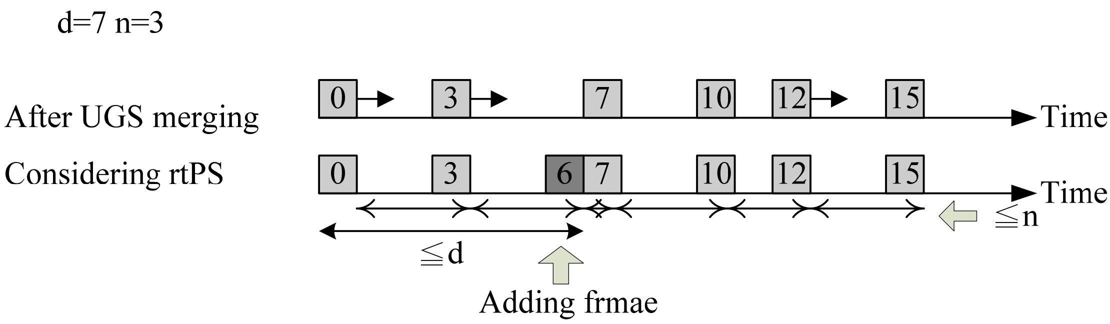

Figure 3.10 An example of transmission merging considering rtPS connections

(3) nrtPS and BE connections merging

Because the nrtPS and BE connections are non-real time, the latency requirement is not considered, and we also consider the minimum reserved rate of nrtPS when calculating the maximum sleep interval. So the nrtPS and BE are transmitted with the rtPS and UGS connections.

3.3.2 Scheduling Strategy

After merging transmissions for PS nodes, we have to allocate BW to the PS MSSs and non-PS MSSs of the BS. For the BS, the first step is allocating BW to the UGS connections and control packets. After that, the residual slots will be divided into uplink and downlink part. The BS will do uplink grant scheduling by the given uplink slots. After uplink grant scheduling, if there are residual BW, it will be released to the downlink part. Then, the BS does the downlink outbound scheduling. After finishing scheduling, the BS sends UL map and DL map with the BW allocation and the next awake-frame of PS MSS. The Step is shown in the Figure 3.11.

After receiving the UL and DL map, the MSS will get itself uplink slots, and then it will do the uplink outbound scheduling.

Step 1.	Satisfy the UGS requirement and control packet
Step 2.	Divide the residual bandwidth into uplink and downlink part
Step 3.	Do uplink grant scheduling and downlink outbound scheduling of BS
Step 4.	Sending UL/DL map
Step 5.	MSS does the uplink out bound scheduling

Figure 3.11 The steps of map generation

Step 1. Satisfy the UGS requirement and control packet

In this step, the UGS allocation of non-PS MSSs is according the UGS duration. Each UGS connection will be set a backoff, and when the backoff is zero, the BS will allocate BW to the UGS connection. Additionally, for PS MSSs, the UGS allocation is decided by transmission merging mechanism.

Step 2. Divide the residual bandwidth into uplink and downlink part

In this step, we decide the uplink and downlink BW by using the total minimum reserved rate of rtPS and nrtPS connections, as shown in (3.6). For the PS node, the r_{min} and r_{max} will be multiplied by the awake interval because the PS MSS have to merge transmission of those frames into an awake-frame.

$$
\begin{aligned}
BW_{available} &= BW_{total} - BW_{UGS} - BW_{control} \\
BW_{uplink} &= \frac{\sum_{\text{rtPS and nrtPS connections of uplink}} r_{min}}{\sum_{\text{rtPS and nrtPS connections of uplink}} r_{min} + \sum_{\text{rtPS and nrtPS connections of downplink}} r_{min}} \\
BW_{downlink} &= \frac{\sum_{\text{rtPS and nrtPS connections of downplink}} r_{min}}{\sum_{\text{rtPS and nrtPS connections of uplink}} r_{min} + \sum_{\text{rtPS and nrtPS connections of downplink}} r_{min}}
\end{aligned}
\tag{3.6}
$$

Step 3. Do uplink grant scheduling and downlink outbound scheduling of BS

When doing scheduling, only non-PS MSS and awake PS MSS will be scheduled. The main idea for the scheduler is to satisfy the minimum requirement of rtPS, and nrtPS firstly, and allocate residual BW fairly to each item. In order to allocate the residual BW fairly, it shares the BW by using the following proportion according to the minimum reserved rate, maximum sustained rate, used rate, and queue size. The used rate is the actually transmission rate before the now scheduled frame. For each item i, there is V_i, where

$$V_i = \mathrm{Min}(r_{max} - r_{min}, r_{max} - r_{used})^+ \quad (3.7)$$

The allocation for the item i is

$$\mathrm{Allocation}_i = \mathrm{Residual\_Bandwidth} \times \frac{V_i}{\sum V_i} \quad (3.8)$$

For the uplink grant scheduler, the item means a MSS. For each MSS, the V_i is calculated by the summation of rtPS/nrtPS minimum reserved rate, the summation of rtPS/nrtPS/BE maximum sustained rate, and the summation of rtPS/nrtPS/BE used rate. For the downlink outbound scheduler, the item means a downlink connection of the BS. The V_i is calculated by minimum reserved rate, maximum sustained rate, used rate of each downlink connection.

After deciding the allocation, the scheduler will check the queue size. If the queue is smaller then the allocation, the residual BW will be released and shared by other connection with non-empty queue using the same proportion. For the PS node, the V_i will be multiplied by the awake interval because the PS MSS have to merge transmission of those frames into a awake-frame.

The example is shown as Figure 3.12. The parameter is shown in Table 3.3. In the beginning, every queue has packets where ID_i means the queue which ID is i. In the first round, the allocation for the ID_1 and the ID_5 is not enough, and the ID_2, and the ID_3 queue requirement is satisfied. The ID_4 can't get any slots because the used rate is larger than the maximum rate. In the second round, there are 15 available slots which are released by ID_2, and ID_3. After allocation, the requirement of ID_5 is satisfied. The ID_1 queue requirement is left 5 slots.

Step 4. Sending map

The step 4 is to sum up total allocated slots of uplink and downlink of the each MSS, and establish the map IE for the each MSS.

Step 5. MSS does the uplink out bound scheduling

The strategy is the same as step 3. For the uplink outbound scheduler of each MSS, the item means a uplink connection of the MSS. The V_i is calculated by minimum reserved rate, maximum sustained rate, used rate of each uplink connection of the MSS.

Table 3.3 The scheduling example parameter

ID	Type	Queue size (slots)	r_{used} (bytes/sec)	r_{min} (bytes/sec)	r_{max} (bytes/sec)
1	rtPS	30	70K	50K	100K
2	rtPS	40	80K	100K	200K
3	nrtPS	30	100K	100K	150K
4	nrtPS	5	120K	50K	100K
5	BE	15	130K	0K	150K

1st Round: Available slots = 100 $\sum V_i=200$

$V_1=\text{Min}(100-50, 100-70)^+ = 30$ => $\text{Allocation}_1=100*30/200=15$

$V_2=\text{Min}(200-100, 200-80)^+ = 100$ => $\text{Allocation}_2=100*100/200=50$ (Release 10 slots)

$V_3=\text{Min}(150-100, 150-100)^+ = 50$ => $\text{Allocation}_3=100*50/200=25$ (Release 5 slots)

$V_4=\text{Min}(100-70, 100-120)^+ = 0$ => $\text{Allocation}_4=0$

$V_5=\text{Min}(150-100,150-130)^+ = 20$ => $\text{Allocation}_5=100*20/200=10$

2nd Round: Available slots = 15 $\sum V_i=50$

$V_1=\text{Min}(100-50, 100-70)^+ = 30$ => $\text{Allocation}_1=15*30/50=10$

$V_2=0$ => $\text{Allocation}_2=0$

$V_3=0$ => $\text{Allocation}_3=0$

$V_4=\text{Min}(100-70, 100-120)^+ = 0$ => $\text{Allocation}_4=0$

$V_5=\text{Min}(150-100,150-130)^+ = 20$ => $\text{Allocation}_5=15*20/50=5$

Figure 3.12 The example of scheduling

3.4 Comparison with AS

In this subsection, we compare the proposed algorithm with AS. In our algorithm, we consider the longer sleep interval, so we can reduce the power consumption of components' on/off. Otherwise, our algorithm can get better performance in merging. As shown in Figure 3.13, the UGS_1 has smallest latency, so AS allocates BW to it firstly. Because the transmission of UGS_1 is already allocated in the 8th frame, it can't be merged with the transmission of UGS_3 in 9th frame. However, in our proposed algorithm, the transmission of UGS_1 will be allocated in 9th frame. Moreover, AS assumes that the connections are known in advance, so it can't add new connections in the run-time.

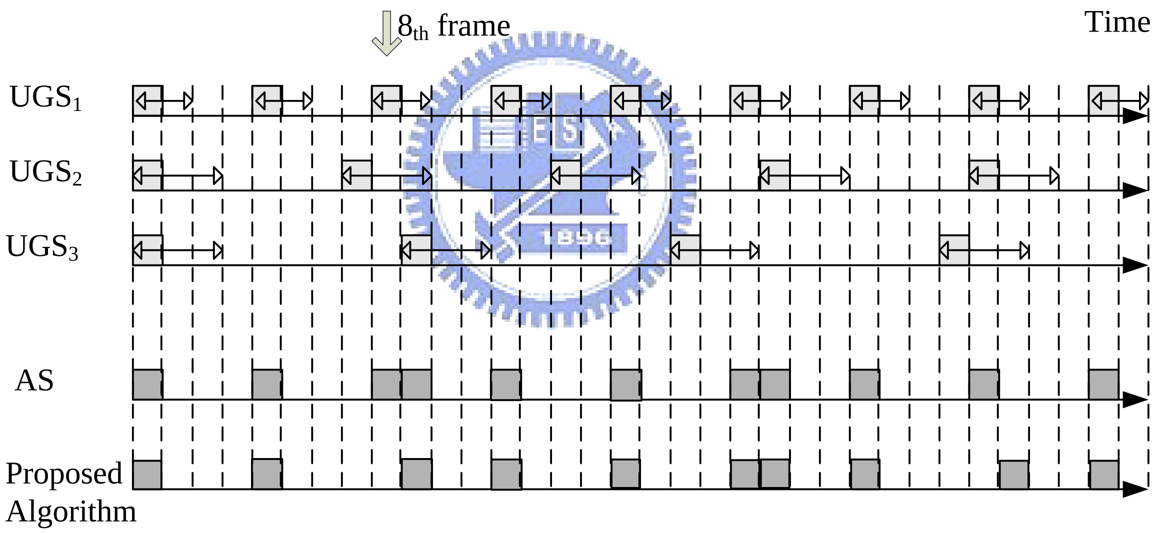

Figure 3.13 An example for comparison between proposed algorithm and AS

Chaper 4 Performance Evaluation

In this chapter, our proposed algorithm is evaluated via simulation.

4.1 Simulation Environment

We assume the BW is 8 Mbps. The duration of a frame is 10 msec, and a time slot is 0.1 msec, respectively. DL-MAP and UL-MAP occupy two slots. All packet arrival occurs at the beginning of a frame. The parameter settings in the simulation are summarized in Table 4.1.

Table 4.1 Parameter settings of simulation

Parameter	Value
Simulation Time	10s
Total bandwidth	8Mbps
Frame duration	10 ms
One Slot duration	0.1ms
Map duration	2 slots
Bytes per slot	100 bytes

4.2 Simulation Results

In this section, we evaluate the scheduling algorithm and show it can allocate BW fairly. We also evaluate the power saving machine and show the performance.

4.2.1 Scenario One

In this scenario, we show the maximum sleep interval's effect in this system, and find the suitable maximum awake interval for the PS node. We also show the proposed formula is reasonable.

4.2.1.1 Simulation Environment

There are two kinds of MSS in this simulation. One is PS MSS. Others are non-PS MSS. The parameter settings of each connection are in Table4.3. We simulate the offline and online merging solutions and show that we choose the proper maximum awake interval for different traffic load.

Table 4.2 Parameter settings of scenario one

	Type	Grant bytes (KB)	Grant Duration (Frames)	Maximum. Latency (Frames)	R_{avg} (KB/sec)	R_{min} (KB/sec)	R_{max} (KB/sec)
PS MSS	UGS	1	10	3			
	UGS	1	12	3			
	rtPS			10	50	40	60
	nrtPS				50	40	60
Non-PS MSS	rtPS				50	40	60
	nrtPS				50	40	60

4.2.1.2 Simulation Result

Firstly, we show the offline merging solution. All the traffic is known in advance. The rtPS traffic incoming duration is calculated by average rate and packet size. There is only one PS MSS in this simulation. We simulate by different maximum awake interval and aggregate the allocated slots in each frame. We show the outcome under different maximum awake interval. The result is shown in Figure 4.1 and Figure 4.2.

The Figure 4.1 is the allocated slots when maximum awake interval is 4. The Figure 4.2 is the allocated slots when maximum awake interval is 8. They show that the proposed merging mechanism can concentrate the allocation. When the maximum awake interval is larger, the allocation is more concentrated. However, when the allocation is more concentrated, it will occupy more slots in a frame and maybe affect other MSS. As shown in the Figure 4.6, the total slot number is 100. In 0.18s, the allocated slots are 80. It means there are only 20 slots for other MSS in that frame.

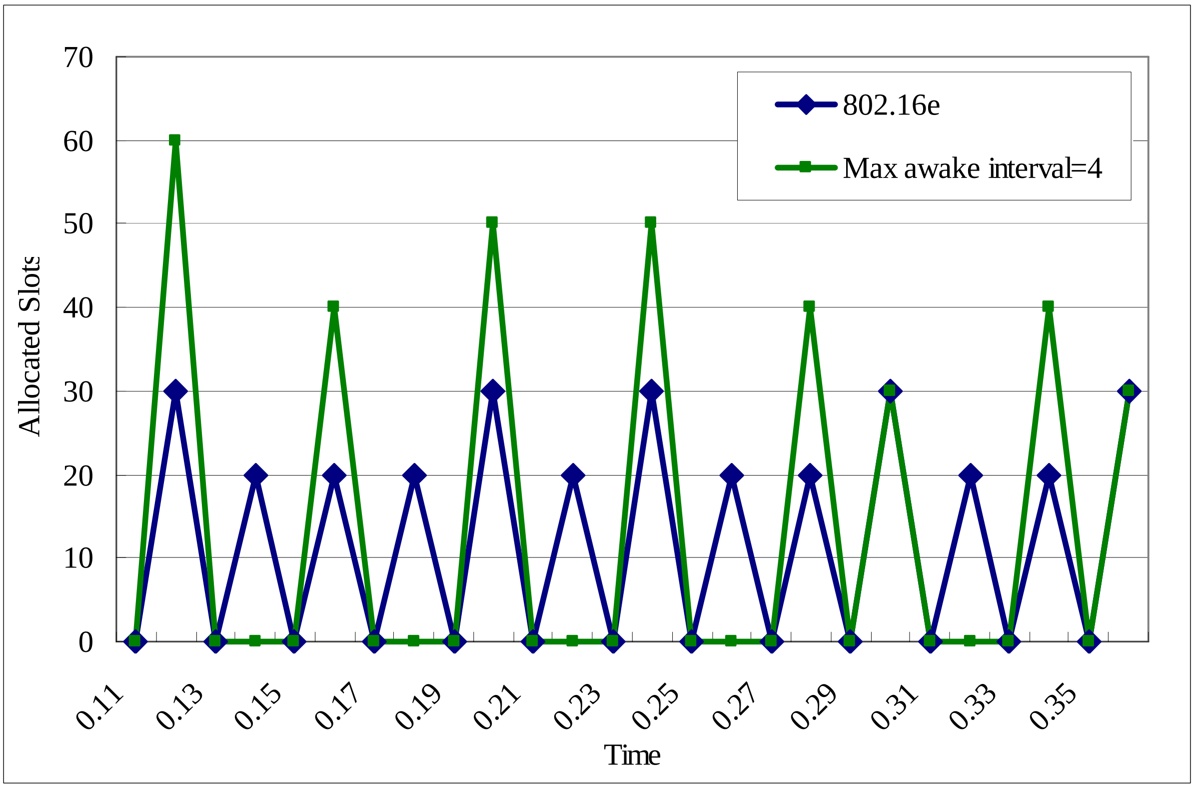

Figure 4.1 The allocated slots with maximum awake interval = 4 in scenario one

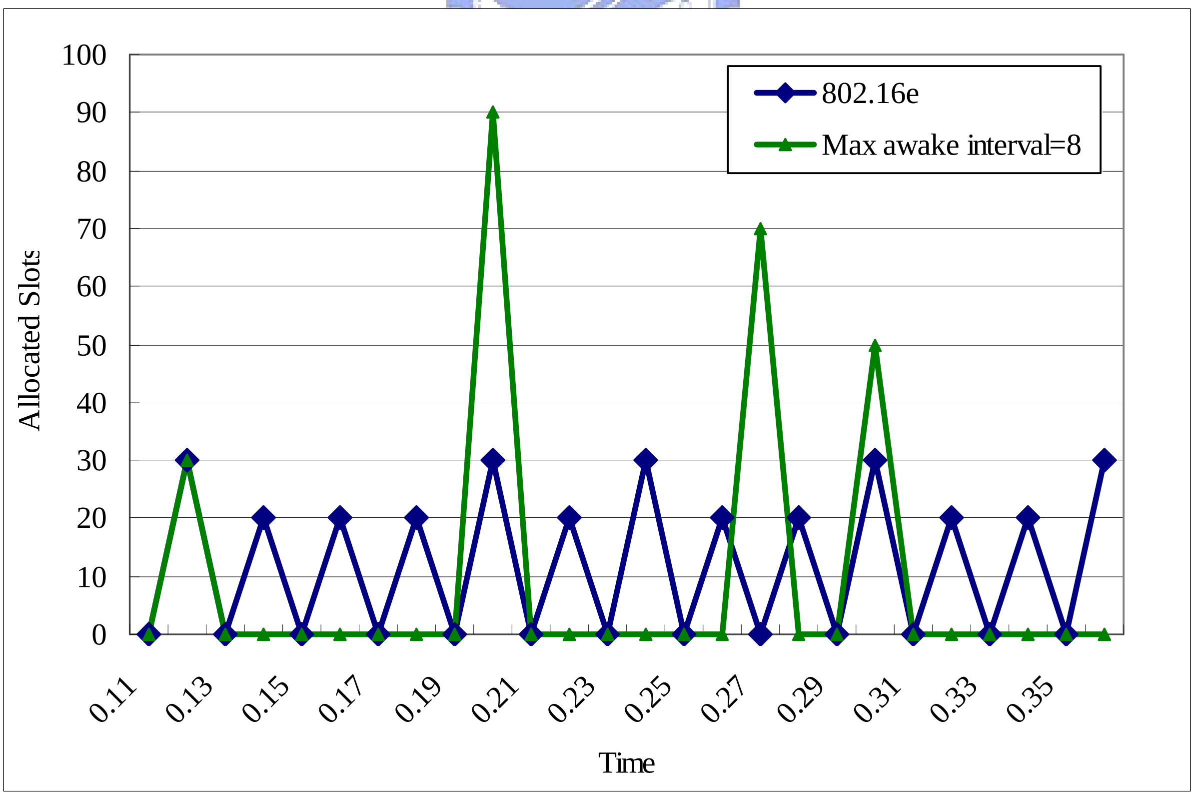

Figure 4.2 The allocated slots with maximum awake interval = 8 in scenario one

Secondly, we set that all the traffic is not known in advance and the traffic is generated by poison processing. We set the simulation three conditions, i.e., four other MSSs, six other MSSs, and eight other MSSs, to represent different traffic load. For four other MSSs, the curve is "MSS = 5". For six other MSSs, the curve is "MSS = 7". For eight other MSSs, the curve is "MSS = 9". We simulate the delay and dropped packets of PS mode MSS with different maximum awake interval and show that we choose the proper maximum awake interval for different traffic load. The results are shown in Figure 4.3 to Figure 4.5.

The Figure 4.3 and Figure 4.4 are the uplink and downlink delay of the PS MSS. When the maximum awake interval is getting higher, the PS MSS average awake interval will be longer. So, the delay of the PS MSS will be larger. Otherwise, when the node number is larger, the delay will be larger, too. But when the delay is increasing, it means some packets are dropped. So we observe the dropped packets number to choose the proper maximum awake interval for different traffic load. It is shown in Figure 4.5.

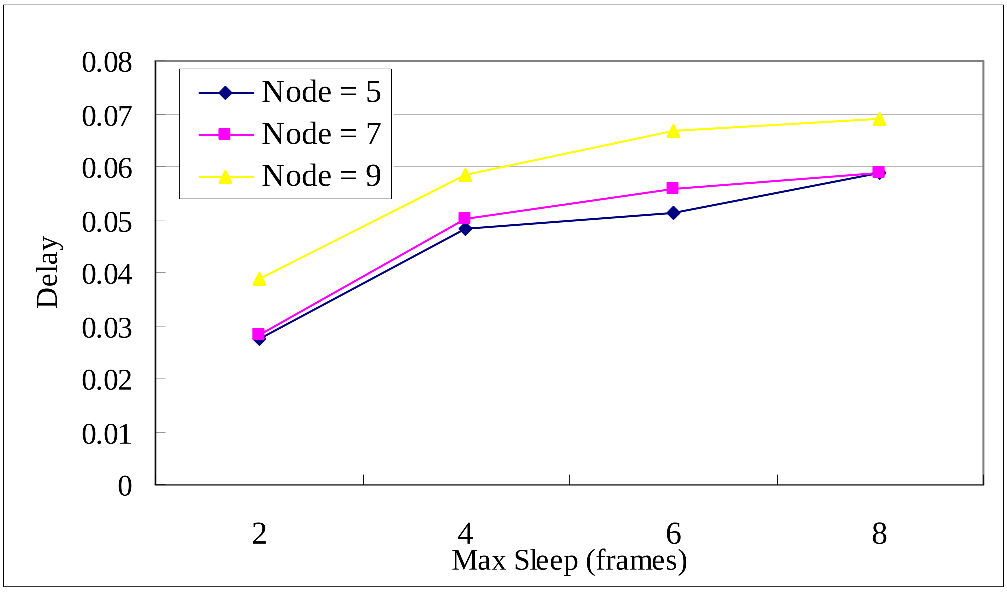

Figure 4.3 The uplink delay for different traffic load of scenario one

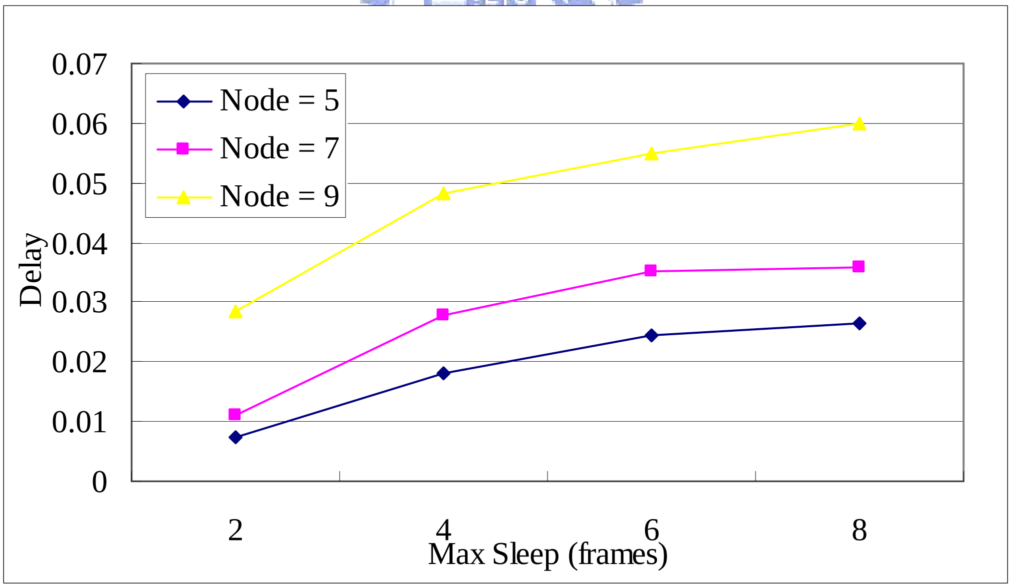

Figure 4.4 The downlink delay for different traffic load of scenario one

As shown in the Figure 4.5, it shows the dropped packets number of three kind of traffic load. For Node = 9, when the maximum awake interval exceeds 2, the system starts to drop packets. For Node = 7, when the maximum awake interval exceeds 4, the system starts to drop packets. For Node = 5, when the maximum awake interval exceeds 6, the system starts to drop packets. So, if we set the proper maximum awake interval for different traffic load, the delay can be bound by the latency requirements and the system avoids dropping packets. In our proposed formula, we calculate maximum awake interval 2 for Node =9, 4 for Node =Node = 7, and 6 for Node = 5. It fits with the simulation.

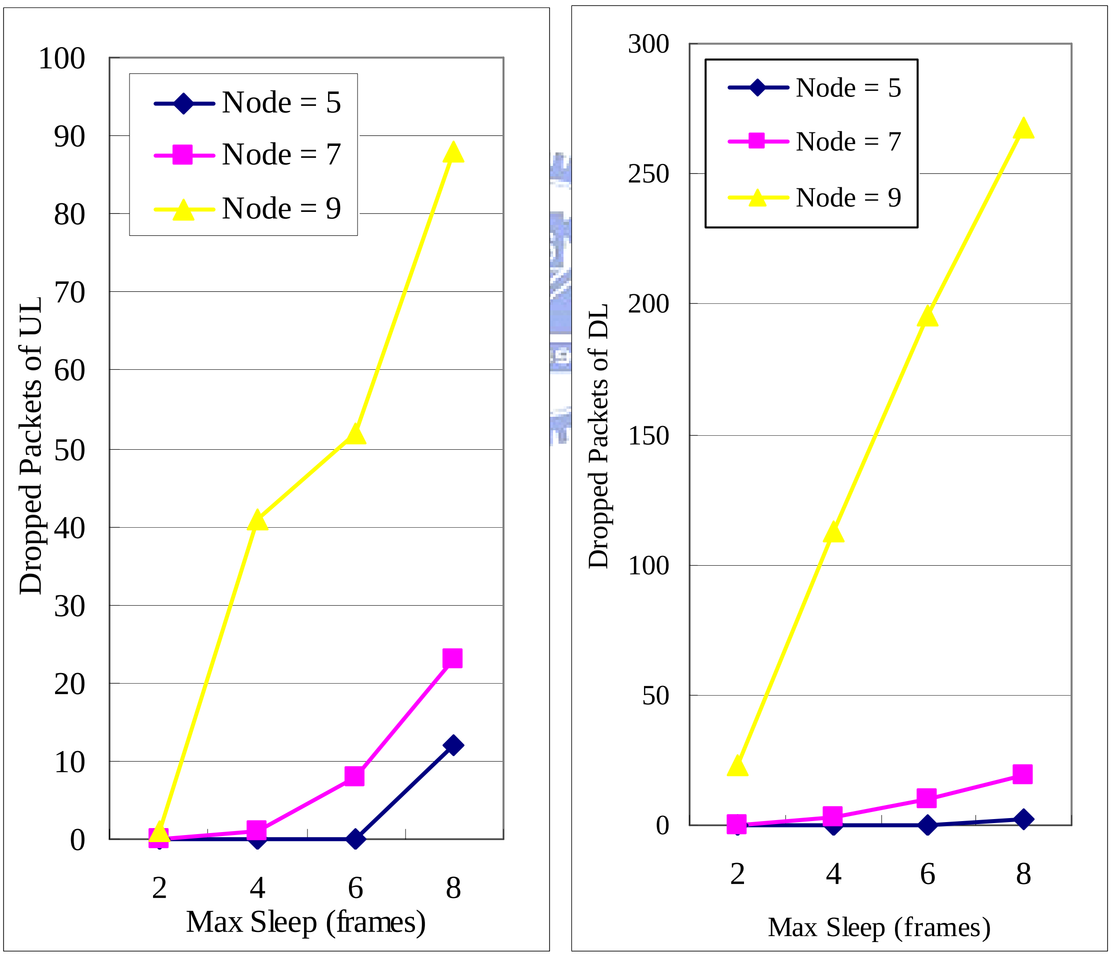

Figure 4.5 The dropped packets number for different traffic load of scenario one

4.2.2 Scenario Two

In this scenario, the simulation shows the fairness of the scheduling. It observes the fairness, delay, and drop rate when raising the traffic load. Otherwise, we show the power saving performance and the movement effect.

4.2.2.1 Simulation Environment

The parameter settings of each MSS are in Table 4.3. Each MSS has one rtPS connection, one nrtPS connection, and one BE connection. The D is a variable. It is supposed that the minimum reserved rate and maximum sustained rate fluctuate about ±20 percent over the average traffic rate. For the PS MSS, we will give other three UGS connections. The parameter settings of UGS the connections are in Table 4.4.

Table 4.3 Parameter settings of scenario two

Type	Average Rate (KB/sec)	R_{min} (KB/sec)	R_{max} (KB/sec)	maximum Latency (frames)
rtPS	20*D	16*D	24*D	10
nrtPS	20*D	16*D	24*D	
BE	20*D		24*D	

Table 4.4 UGS parameter settings of scenario two

Type	Direction	Grant bytes (KB)	Grant Duration (Frames)	Maximum Latency (Frames)	R_{min} (KB/sec)	R_{max} (KB/sec)
UGS	DL	1	8	3		
UGS	DL	1	10	3		
UGS	UL	1	7	2		

Here, we adjust the fairness of different classes according the formula 4.1.

$$FAIR_{a,b} = \left| \frac{Th_a}{S_a} - \frac{Th_b}{S_b} \right| \quad (4.1)$$

The "a" and "b" represent two different classes. The "Th_a" and "Th_b" represent the throughputs of class "a" and "b" by calculating the transmitted packets size. The "S_a" and "S_b" represent total generated packets size by the source. When the "$FAIR_{a,b}$" is smaller, it means the two classes is fairer.

Because the actual power consumption of on/off is unknown. So we show the used power by calculating the awake-frames, and define power consumption as following formula 4.2.

$$\text{Power consumption} = \frac{\text{Awake frame}}{\text{Total frame}} \quad (4.2)$$

4.2.2.2 Simulation Result

Firstly, we show the fairness of the scheduling. We set eight MSSs in the system, give different D to increase the traffic, and calculate the results. Fairness, delay and drop rate are investigated in performance study. Here, we compare the rtPS V.S nrtPS and rtPS V.S BE.

The Figure 4.6 and 4.7 shows the fairness of the rtPS V.S nrtPS and rtPS V.S BE. The X axis is the summation of all connections' average rate. As shown in Figure 4.6, when the average rate is less than 1440 KB/sec, it shows good fairness of rtPS V.S nrtPS with beyond 0.05. The BS can satisfy both the minimum reserved rate of rtPS and nrtPS. But, when the total average rate exceeds 1440 KB/sec, the BS can only satisfy the minimum reserved rate of rtPS. The fairness value increases because the BS serves the rtPS minimum requirement firstly. Until the average rate exceeds to 2400Kbytes, the BW can't serve the rtPS minimum reserved rate and then, the Fairness value decreases.

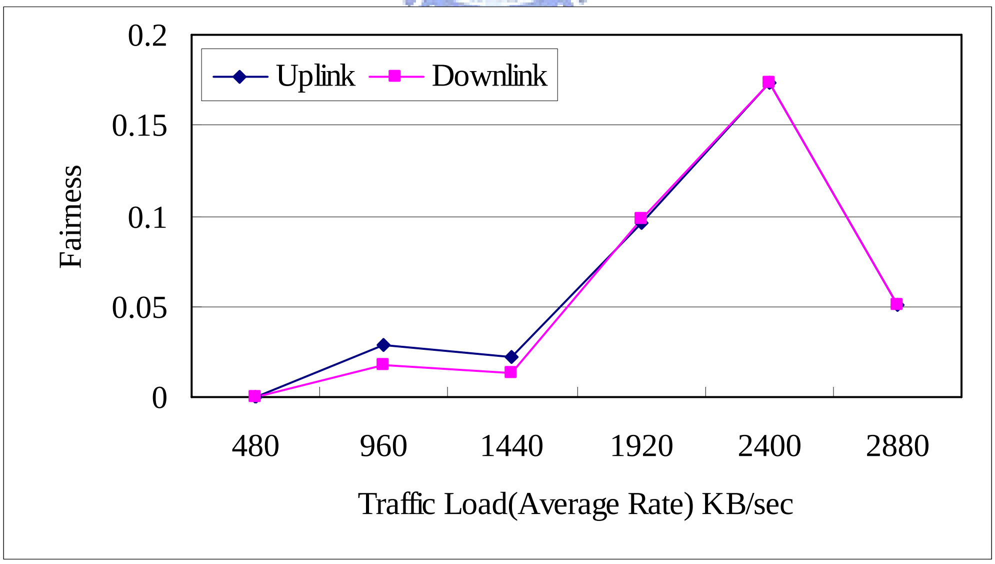

Figure 4.6 Fairness of rtPS V.S nrtPS in scenario two

As shown in Figure 4.7, When the total average rate is less then 960 KB/sec, the Fairness value is less then 0.4. The BE still can get BW. However, when the total average rate exceeds to 960 KB/sec, the fairness is getting increasing. It’s because it have to satisfy the minimum reserved rate of rtPS and nrtPS firstly and there are no enough BW for BE. When the total average rate exceeds to1440 KB/sec. the BE can’t get any BW because the capacity only can serve the rtPS and nrtPS service’s minimum reserved rate.

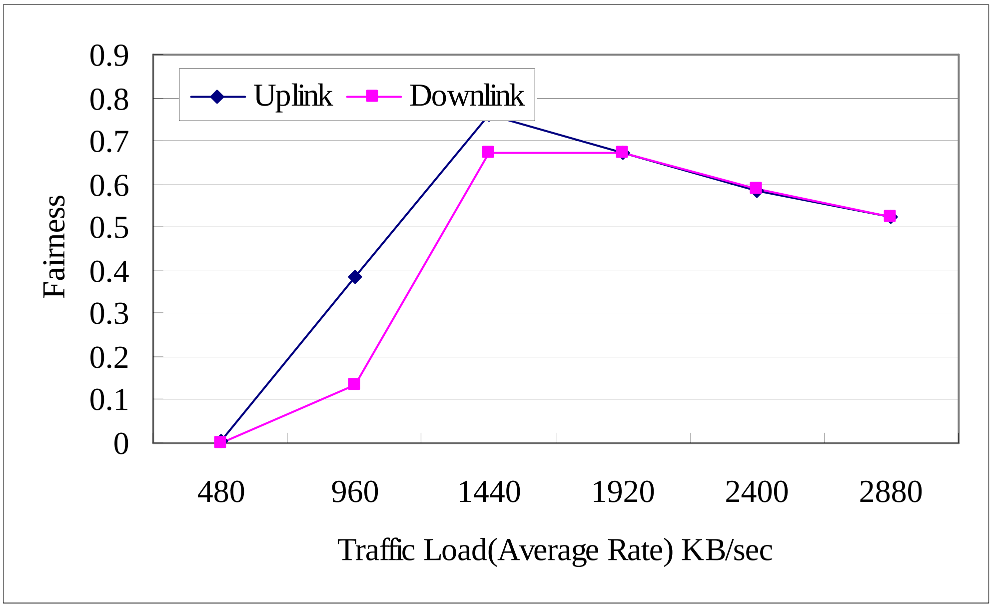

Figure 4.7 Fairness of rtPS V.S BE in scenario two

Figure 4.8 is the delay and Figure 4.9 is the drop rates of rtPS. They can observe that when the traffic load exceeds 960Kbytes, the delay and drop rate of rtPS is getting higher.

From Figure 4.6 to Figure 4.9, we can observe that the performance of downlink is better then uplink. It is because the uplink information is older then downlink. When the packets of uplink arrive, it will be known and updated by BS when uploading date. It must be scheduled until the next frame. So, the performance of uplink is worse then downlink.

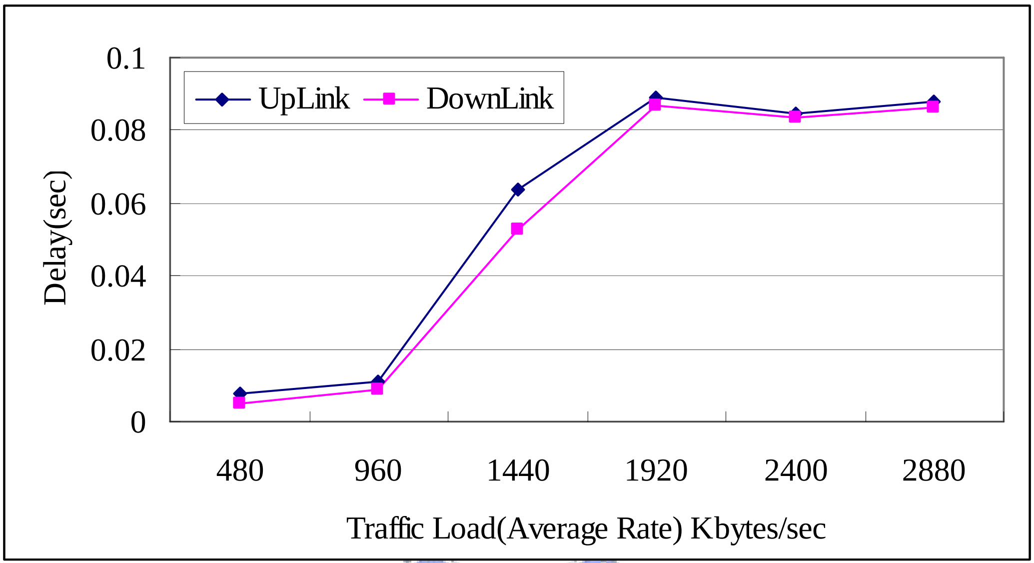

Figure 4.8 Delay of rtPS in scenario two

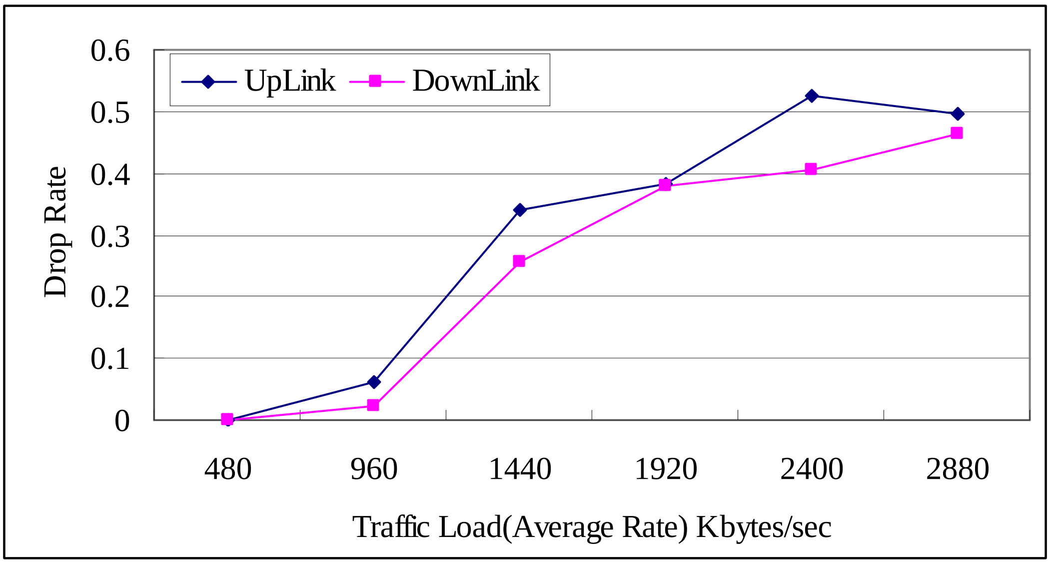

Figure 4.9 Drop rates of rtPS in scenario two

Secondly, we show the power saving performance and the movement effect. There are six MSSs and the D is set 3. One of the six MSSs will leave the system when t = 5. One of the six MSSs is PS node with three UGS connections and enters the PS mode in the beginning. We simulate the power saving performance with/without transmission merging mechanism by comparing power consumption and delay.

Figure 4.10 is the accumulated frame number. Figure 4.11 is the power consumption. We can observe that PS MSS in the 802.16e sends or receives packets more then 80 frames in 1 sec. It means the PS MSS have to work more then 80 frames in 1 sec. If using the propose algorithm, there are no more then 40 frames for working by merging the transmission.

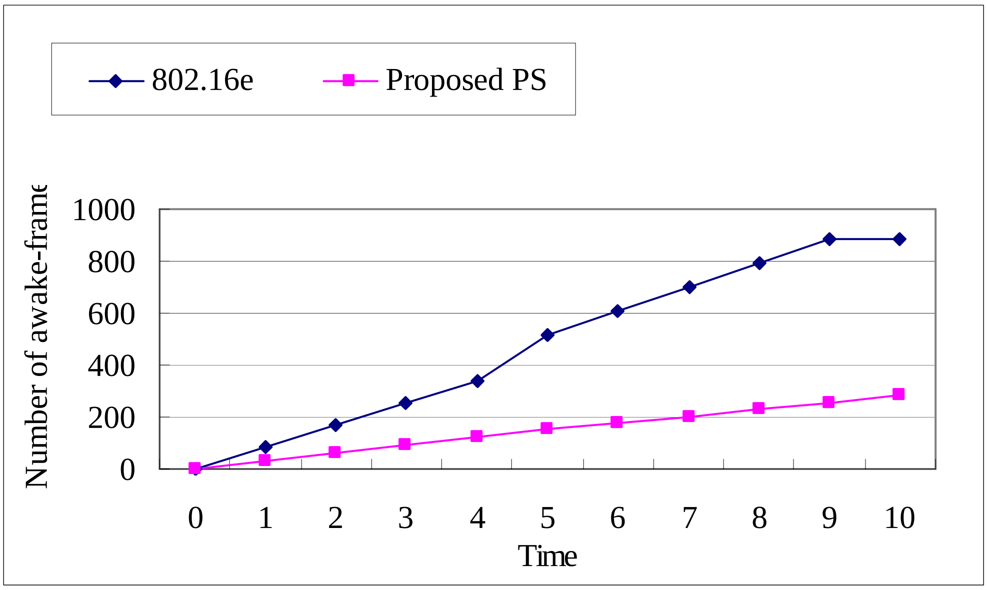

Figure 4.10 Accumulated working frame of PS MSS in scenario two

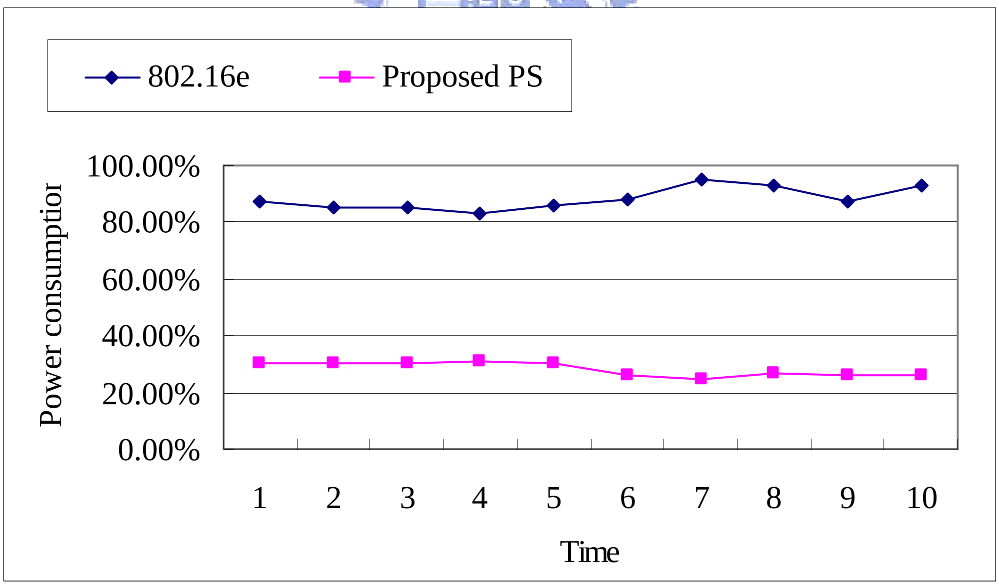

Figure 4.11 Power consumption of PS MSS in scenario two

As shown in Figure 4.12 and Figure 4.13, the delay of uplink and downlink of PS MSS using the proposed algorithm will be larger then the original 802.16e and the delay is still bounded by the requirement 10 ms. Otherwise after 5 sec, the traffic load is lighter, so the calculated maximum awake interval is longer to lead the delay of UL/DL larger.

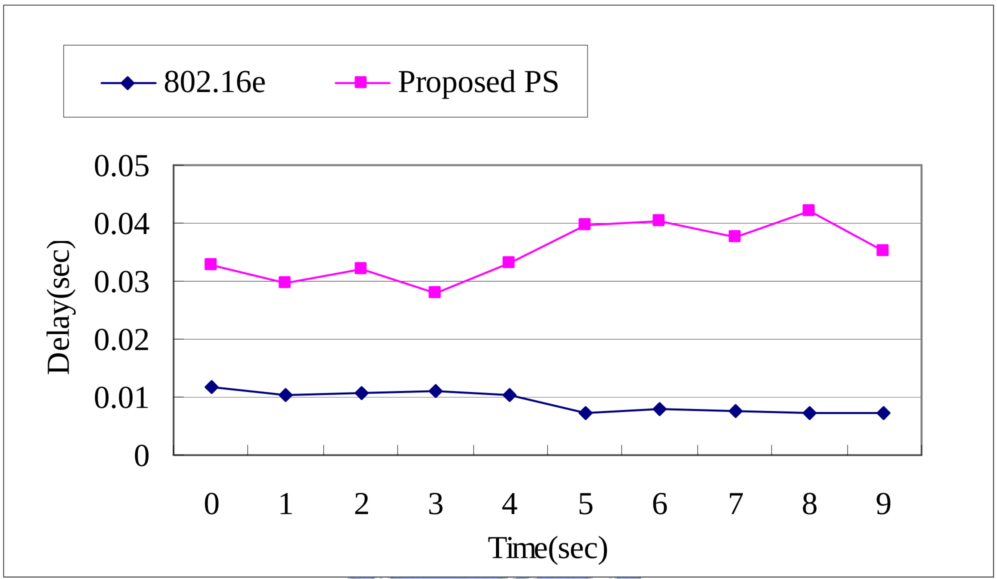

Figure 4.12 Uplink delay of PS MSS in scenario two

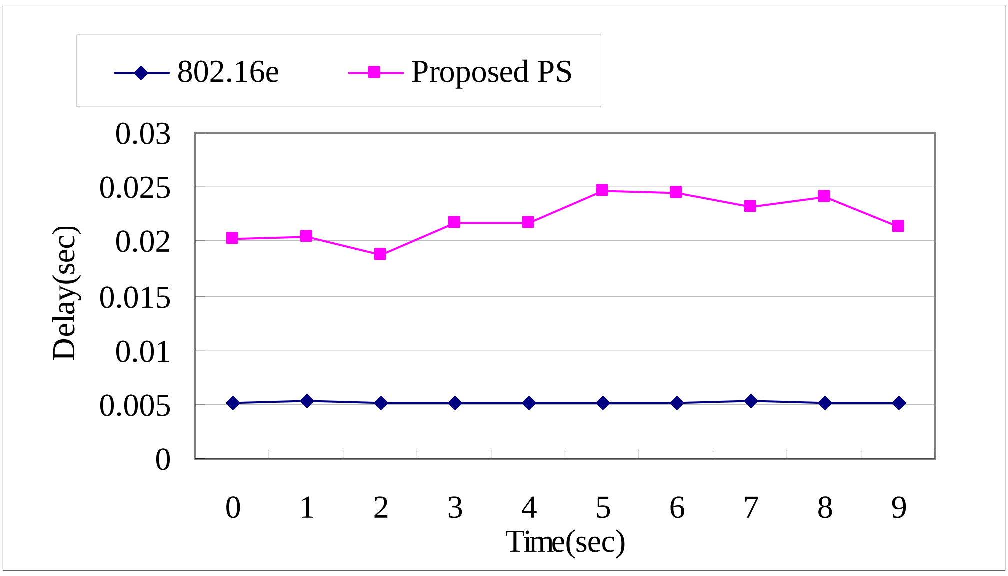

Figure 4.13 Downlink delay of PS MSS in scenario two

Chaper 5 Conclusion and Future Work

In this thesis, we propose a novel power saving framework which includes a transmission merging mechanism and a scheduler for the IEEE 802.16e. The transmission merging mechanism reduces the number of awake-frames by merging the connections transmission time. We consider both real-time and non real-time connections, satisfy the QoS requirement, and get better merging performance then AS. In our simulation, we show our algorithm can get the proper maximum awake interval. Then, we evaluate the fairness under different traffic load. The results show that our scheduling algorithm can achieve fairness under QoS satisfactions. Otherwise, we evaluate the power consumption and delay with/without transmission merging mechanism. The results show that our mechanism can get good power efficiency degree under QoS satisfactions.

In this thesis, we only consider one QoS parameter, i.e., the maximum latency, for the UGS connections. However, the delay jitter is another important QoS parameter to characterize UGS connections' behavior. Thus we will focus on jitter in our following work. However, we will take ertPS connections into consideration.

Reference

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

[2] IEEE 802.16e-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, 2006.

[3] Changhoi Koo, and Yong Woo Chung, "Dynamic Cell Coverage Control for Power Saving in IEEE802.16 Mobile MultihopRelay Systems," IARIA/IEEE ICSNC, pp. 60-60, 2006.

[4] Jeffrey P. Monks, Vaduvur Bharghavan, and Wen-mei W. Hwu, “A Power Controlled Multiple Access Protocol for Wireless Packet Networks,” IEEE INFOCOM, vol.1, pp. 219-218, 2001.

[5] Jen-Mao Huang, “Fair Queuing with Power Saving in Wireless Ad Hoc Network,” master thesis in NCTU, 2006.

[6] Yung-Hsiang Lu, Eui-Young Chung, Tajana Simunic, Luca Benini, and Giovanni De Micheli, “Quantitative Comparison of Power Management Algorithms,” DATE Proceedings, pp. 20-26, 2000.

[7] Jianfeng Chen, Wenhua Jiao, and Hongxi Wang, ”A Service Flow Management Strategy for IEEE 802.16 Broadband Wireless Access Systems in TDD Mode,” IEEE ICC, Vol. 5, pp. 3422-3426, 2005.

[8] Yang Xiao, “Energy Saving Mechanism in the IEEE 802.16e Wireless MAN,” IEEE Communications Letters, pp. 595-597, 2005.

[9] Jun-Bae Seo, Seung-Que Lee, Nam-Hoon Park, Hyong-Woo Lee, and Choong-Ho Cho, “Performance Analysis of Sleep Mode Operation in IEEE802.16e,” IEEE VTC, pp. 1169-1173, 2004.

[10] Yang Xiao, "Performance Analysis of an Energy Saving Mechanism in the IEEE 802.16e Wireless MAN," IEEE CCNC Proceeding, pp. 406-410, 2006.

[11] Yan Zhang, and Masayuki Fujise, "Energy Management in the IEEE 802.16e MAC," IEEE Communication Letters, pp. 311-313, 2006.

[12] Lei Kong, and Danny H.K.Tsang, "Performance Study of Power Saving Classes of Type I and II in IEEE 802.16e," IEEE LCN Proceedings, pp. 20-27, 2006.

[13] You-Lin Chen, and Shiao-Li Tsao, "Energy-Efficient Sleep-Mode Operation for Broadband Wireless Access Systems," IEEE VTC, pp. 1-5, 2006.