The Eighth International Symposium on Operations Research and Its Applications (ISORA'09)
Zhangjiajie, China, September 20–22, 2009

Performance Analysis for Power Saving Class Type II in IEEE 802.16e for Wireless MAN

Shunfu Jin[1] Ranran Li[1] Wuyi Yue[2]

[1]College of Information Science and Engineering, Yanshan University, Qinhuangdao 066004 China
[2]Department of Intelligence and Informatics, Konan University, Kobe 658-8501 Japan

Abstract Effective power saving mechanism is the most significant issue for extending the lifetime of mobile stations in wireless metropolitan area network (MAN). In this paper, we propose an effective method to analyze the system performance of the power saving class type II in IEEE 802.16e. We present a queueing model with two kinds of vacation mechanisms to capture the working principle of the networks with the power saving class type II. We also propose measure methods in terms of the energy saving ratio, the handover ratio and the average response time, and give the expressions for these performance measures. Moreover, considering both the energy saving ratio and the average response time, we develop a cost function to determine the optimal length of the sleep window for minimizing the total system cost. Numerical illustrations are provided to demonstrate the relationship between the system performance and the configuration parameters. The research in this paper provides a theoretical basis for improvement of the power saving class type II, and has potential application when solving other energy conserving related problems in wireless mobile networks.

Keywords Mobile broadband metropolitan area networks; Power saving class type II; Multiple vacation queueing system; Performance analysis and evaluation

1 Introduction

Energy is a scarce resource in wireless metropolitan area network (MAN), it is critical to design energy efficient techniques to extend the lifetime of the battery in the Mobile Station (MS). To support battery powered mobile Broadband Wireless Access (BWA) devices efficiently, the IEEE 802.16e offers three kinds of sleep modes called power saving classes of type I, type II and type III, respectively. Power saving class type II is mainly used for the Unsolicited Grant Service (UGS) and the Real Time Variable Rate (RT-VR) traffic.

Several authors have shown interest in the performance of the sleep mode operation, either in the case of the IEEE 802.16e or other technologies.

Xiao was the first to study the energy saving efficiency of the sleep mode in IEEE 802.16e, and the energy consumption and the mean delay for power saving class type I by focusing on a sleep interval was obtained in [1].

B. Lee and H.Lee proposed a cumulative-TIM (Traffic Indication Message) method in order to improve the energy efficiency for the power saving class type I in [2]. By the method of simulation, the MS selects the length of the sleep window for determining the trade-off function between the energy saving efficiency and the data delay.

Han and Choi modeled the Base Station (BS) buffer as a continuous-time finite-capacity queue with a Poisson arrival process and a deterministic service time, and derived the expressions for the average packet delay and the average energy consumption with a semi-Markov chain in [3].

Lee and Cho investigated the performance of the power saving class type II in VoIP traffic, addressed the problem of allocation representation of VoIP packets, and proposed an efficient uplink mapping scheme by using a simulation method in [4].

However, most of the works mentioned above were based on continuous-time stochastic process, and there is no comprehensive theoretical analysis on the power saving class type II until now. In order to improve the energy saving efficiency of type II and evaluate the system performance efficiently, novel analytical methods must be provided. On the other hand, it is indicated that it would be more accurate and efficient using discrete time models than continuous time counterparts when analyzing and designing digital transmitting systems [5], [6]. Moreover, vacation queue model is naturally more suitable for the research of the sleep mode in power saving schemes.

In this paper, we propose an effective analysis method to evaluate the system performance of the power saving class type II in IEEE 802.16e for wireless MAN. Taking into account the memoryless character of the data frame arrival and digital nature in the power saving class type II applied in UGS and RT-VR traffic, and considering the fact that some data frames could be transmitted during the listening state, in this paper, we model the system as a Geom/G/1 queue model with two kinds of multiple vacation mechanisms. One kind is the normal vacation mechanism representing the sleep state, the other kind is a special vacation mechanism representing the listen sate, in which a limited number of data frames can be transmitted. By using an embedded Markov chain method and the boundary state variable theory, we give the performance measures of the system, and also present numerical results to show the system performance with system parameters.

2 Principle for Power Saving Class Type II and System Model

For all the sleep modes in the power saving schemes I, II and III in IEEE 802.16e, the MS operates in three state: awake state, sleep state and listen state. The lengths for the system being in the sleep state and the listen state are controlled by sleep window and listen window, respectively. In the awake state, the MS or the BS can transmit data frames normally. In the sleep state, the MS conducts a pre-negotiated period of absence from the serving BS air interface. In the listen state, the MS senses the channel all the time to see if there is data frame to be transmitted.

Different any from other two power saving mechanisms of types I and III, the lengths of the sleep window and the listen window in the power saving class type II are fixed, and a certain number of data frames can be transmitted in the listen state. If all the data frames buffered have been transmitted during the listen state, the system will return back to the sleep state after the listen period is over, otherwise, the system will enter into the awake state to transmit the remainder data frames. It is illustrated in IEEE 802.16e that the MS will awake up at any time when a uplink traffic arrives. It means that the delay of uplink traffic is independent of the sleep mechanism, so we focus on the downlink traffic only in this paper.

Regard the sleep state as one vacation period V_S with the length of T_S, the whole listen state as another special vacation period V_L with the length of T_L. Note that during the vacation period V_L, some data frames can be transmitted. The time period for transmitting data frames in the listen state (in the vacation period V_L) is seen as one busy period B_L with the length of T_{BL}. The interval time for transmitting data frames normally in awake state is regarded as another busy periods B_A with the length of T_{BA}. The busy period B of the system is composed of one or more busy period B_L in a listen state and only one busy period B_A in an awake state. Therefore, we can build a Geom/G/1 queueing model with two kinds of multiple vacation mechanisms. One kind is the normal vacation mechanism representing the sleep state, the other is a special vacation mechanism representing the listen sate.

The transmission time of a data frame is assumed to be independent and identically distributed random variable denoted by S. The probability distribution, the Probability Generating Function (P.G.F.) and the average of S are given as follows:

$$s_k = P\{S = k\}, \quad k \geq 1, \quad S(z) = \sum_{k=1}^{\infty} s_k z^k, \quad E[S] = \sum_{k=1}^{\infty} k s_k.$$

Suppose that the maximal number of data frames can be transmitted within a listen period is d. To simplify the analysis procedure, we neglect the data frame arrivals during the short listen window, so we have $d = T_L/E[S]$.

Taking into account the memoryless nature of users initiated data frame arrival, we can suppose the arrival process to follow a Bernoulli distribution with arrival ratio p $(0 < p < 1)$. We choose the embedded Markov points at the end of slots where the date frame transmissions completed. The sufficient and necessary condition for this embedded Markov chain to be positive recurrent is $\rho = pE[S] < 1$, where ρ is the system load.

3 Busy Period

In the power saving class type II, there are two busy periods, namely, the busy period B_L in the listen state and the busy period B_A in the awake state as we presented in Section 2.

3.1 Busy Period in Listen State

In this system model, the time axis is divided into segments of equal length called slots. We assume that the arrival and the departure of data frames occur only at the boundary of a slot. The state of the system is defined by the number of data frames at the embedded Markov points. It is assumed that data frames are transmitted according to a First-Come First-Served (FCFS) discipline.

Under the condition that there is at least one data frame arrival in the sleep state, we assume the number of data frames transmitted during the listen state to be Q_{BL}. Then the average $E[Q_{BL}]$ of Q_{BL} can be obtained as follows:

$$E[Q_{BL}] = \frac{\sum_{j=1}^{d-1} j \binom{T_s}{j} p^j \bar{p}^{T_S - j} + \sum_{j=d}^{T_S} d \binom{T_s}{j} p^j \bar{p}^{T_S - j}}{1 - \bar{p}^{T_S}} \tag{1}$$

where $\bar{p} = 1 - p$, T_S is the length of the vacation period V_S, j is the number of data frames arrived in the sleep state. Therefore, the average $E[T_{BL}]$ for the busy period length T_{BL} in the listen state can be given by

$$E[T_{BL}] = \frac{\sum_{j=1}^{d-1} j \binom{T_s}{j} p^j \bar{p}^{T_S - j} + \sum_{j=d}^{T_S} d \binom{T_s}{j} p^j \bar{p}^{T_S - j}}{1 - \bar{p}^{T_S}} E[S]. \tag{2}$$

3.2 Busy Period in Awake State

Under the condition that the number of the data frames arrived in the sleep state exceeds d, after all the previous d data frames are completely transmitted, and also the listen window expired, then the system switches to an awake state to transmit the residual data frames continuously. d is the maximal number of data frames that can be transmitted within a listen period defined in Section 2. Let the number of data frames at the beginning instant of an awake state be Q_{BA}, the average $E[Q_{BA}]$ of Q_{BA} can be obtained as follows:

$$E[Q_{BA}] = \frac{\sum_{j=d+1}^{T_S} j \binom{T_s}{j} p^j \bar{p}^{T_S - j}}{\sum_{j=d+1}^{T_S} \binom{T_s}{j} p^j \bar{p}^{T_S - j}} - d. \tag{3}$$

From [5], we can obtain the average $E[T_{BA}]$ of the busy period length T_{BA} in an awake state as follows:

$$E[T_{BA}] = E[Q_{BA}] \frac{E[S]}{1-\rho} = \left(\frac{\sum_{j=d+1}^{T_S} j \binom{T_s}{j} p^j \bar{p}^{T_S - j}}{\sum_{j=d+1}^{T_S} \binom{T_s}{j} p^j \bar{p}^{T_S - j}} - d \right) \frac{E[S]}{1-\rho}.$$

3.3 Busy Cycle

The busy cycle R is defined as a time period from the instant in which the busy period B of the system completes to the instant in which the next busy period B of the system ends. Let N_{BL} be the number of busy periods B_L in a busy cycle B, the average $E[N_{BL}]$ of N_{BL} is given by

$$E[N_{BL}] = \frac{1}{\sum_{j=d+1}^{T_S} \binom{T_s}{j} p^j \bar{p}^{T_S - j}}. \tag{4}$$

Let T_B be the length of the busy period B of the system, the average $E[T_B]$ of T_B is obtained by

$$\begin{aligned} E[T_B] &= (1-\bar{p}^{T_S})E[N_{BL}]E[T_{BL}]+E[T_{BA}] \\ &= \frac{\rho T_S - \rho E[S] \sum_{j=0}^{d} j \binom{T_s}{j} p^j \bar{p}^{T_S-j} + \rho T_L \sum_{j=0}^{d} \binom{T_s}{j} p^j \bar{p}^{T_S-j} - \rho T_L}{\sum_{j=d+1}^{T_S} \binom{T_s}{j} p^j \bar{p}^{T_S-j}(1-\rho)}. \end{aligned} \tag{5}$$

Let T_R be the length of a busy cycle R, the average $E[T_R]$ of T_R is given as follows:

$$\begin{aligned} E[T_R] &= E[N_{BL}](T_S+T_L)+E[B]E[Q_{BA}] \\ &= \frac{T_S - E[S] \sum_{j=0}^{d} j \binom{T_s}{j} p^j \bar{p}^{T_S-j} + T_L \sum_{j=0}^{d} \binom{T_s}{j} p^j \bar{p}^{T_S-j} - \rho T_L}{\sum_{j=d+1}^{T_S} \binom{T_s}{j} p^j \bar{p}^{T_S-j}(1-\rho)}. \end{aligned} \tag{6}$$

4 Waiting Time

We perform the waiting time analysis in two cases: (1) The waiting time W_L for the data frames transmitted in the listen state. (2) The waiting time W_A for the data frames transmitted in the awake state. In the following subsections, we present how to analyze W_L and W_A, respectively.

4.1 Waiting Time W_L

The waiting time W_L for the data frames transmitted in a listen state can be divided into two parts: (1) The residual time of a sleep window denoted as T_S^+. (2) The time elapsed during listen state denoted as W_{dL}.

The residual time T_S^+ of a sleep window is the period from the instant of a data frame arrived in this sleep window to the end of this sleep window. The average $E[T_S^+]$ of T_S^+ is given by

$$E[T_S^+] = \frac{T_S - 1}{2}. \tag{7}$$

Using the boundary state variable theory presented in [5], we can get the average $E[W_{dL}]$ of W_{dL} as follows:

$$E[W_{dL}] = \frac{\sum_{j=1}^{d-1} j(j-1)\binom{T_s}{j} p^j \bar{p}^{T_S-j} + \sum_{j=d}^{T_S} d(d-1)\binom{T_s}{j} p^j \bar{p}^{T_S-j}}{2\left(\sum_{j=1}^{d-1} j \binom{T_s}{j} p^j \bar{p}^{T_S-j} + \sum_{j=d}^{T_S} d \binom{T_s}{j} p^j \bar{p}^{T_S-j}\right)} E[S]. \tag{8}$$

Combining Eq. (7) and Eq. (8), the average $E[W_L]$ of the waiting time W_L is given by

$$\begin{aligned} E[W_L] &= E[T_S^+]+E[W_{dL}] \\ &= \frac{T_S-1}{2} + \frac{\sum_{j=1}^{d-1} j(j-1)\binom{T_s}{j} p^j \bar{p}^{T_S-j} + \sum_{j=d}^{T_S} d(d-1)\binom{T_s}{j} p^j \bar{p}^{T_S-j}}{2\left(\sum_{j=1}^{d-1} j \binom{T_s}{j} p^j \bar{p}^{T_S-j} + \sum_{j=d}^{T_S} d \binom{T_s}{j} p^j \bar{p}^{T_S-j}\right)} E[S]. \end{aligned} \tag{9}$$

4.2 Waiting Time W_A

The waiting time W_A can be obtained by the sum of two independent random variables, i.e., $W_A = W_{0A} + W_{dA}$, where W_{0A} is the waiting time for the classical Geom/G/1 queue, and W_{dA} is the additional waiting time by the vacations introduced in this system. From [5], we can obtain the P.G.F. $W_{0A}(z)$ and the average $E[W_{0A}]$ of W_{0A} as follows:

$$W_{0A}(z) = \frac{(1-\rho)(1-z)}{(1-z)-\rho(1-G(z))}, \quad E[W_{0A}] = \frac{p}{2(1-\rho)}E[S(S-1)]. \tag{10}$$

Applying the boundary state variable theory, we can get the P.G.F. $W_{dA}(z)$ and the average $E[W_{dA}]$ of W_{dA} as follows:

$$W_{dA}(z) = \frac{p\left(1-Q_{BA}\left(\frac{z-\bar{p}}{p}\right)\right)}{E[Q_{BA}](1-z)}, \quad E[W_{dA}] = \frac{\sum_{j=d+1}^{T_S}(j-d)(j-d-1)\binom{T_s}{j}p^j\bar{p}^{T_S-j}}{2p\sum_{j=d+1}^{T_S}(j-d)\binom{T_s}{j}p^j\bar{p}^{T_S-j}}. \tag{11}$$

The data frames transmitted in the awake state can be also classified into two category: (1) Data frames arrived in the listen state and transmitted in the awake state. The probability for this case is $1-\rho$. Denote the waiting time for this kind of data frames as W_{A1}. (2) Data frames arrived and transmitted both in the awake state. The probability for this case is ρ, where ρ is the system load defined in Section 2. Denote the waiting time for this kind of data frames as W_{A2}.

The data frames arrived in the listen state and transmitted in the awake state will go through a listen period before they being transmitted, so the expression of the waiting time W_{A1} for these data frames is $W_{A1} = T_L + W_{0A} + W_{dA}$, where T_L is the length of a listen window defined in Section 2. For the data frames arrived and transmitted both in the awake state, the waiting time W_{A2} is the sum of W_{0A} and W_{dA}, i.e., $W_{A2} = W_{0A} + W_{dA}$. So the average $E[W_A]$ of the waiting time W_A is then given by

$$E[W_A] = (1-\rho)E[W_{A1}] + \rho E[W_{A2}]. \tag{12}$$

4.3 System Waiting Time W

From the discussions above, we know that data frames will be transmitted either in the listen state, or the awake state. Let P_L be the probability that a data frame is transmitted in the listen state, and P_A be the probability that a data frame is transmitted in the awake state. The average $E[W]$ of the system waiting time W is given as follows:

$$E[W] = P_L E[W_L] + P_A E[W_A]$$

where the expressions for P_L and P_A are given as follows:

$$P_L = \frac{\sum_{j=1}^{d}\binom{T_s}{j}p^j\bar{p}^{T_S-j} + \sum_{j=d+1}^{T_S}\binom{T_s}{j}p^j\bar{p}^{T_S-j}\frac{d}{j}}{1-\bar{p}^{T_S}}, \quad P_A = \frac{\sum_{j=d+1}^{T_S}\binom{T_s}{j}p^j\bar{p}^{T_S-j}\left(1-\frac{d}{j}\right)}{1-\bar{p}^{T_S}}. \tag{13}$$

5 Performance Measures

We define the handover ratio β as the number of switches from the sleep state to the awake state per slot, it is one of the important measures for evaluating the energy consumption. The handover ratio β is given by

$$\beta = \frac{1}{E[T_R]}. \tag{14}$$

The energy saving ratio γ is defined as the amount of energy saved per slot, which is an important performance measure for evaluating the energy saving efficiency for the power saving class type II. The energy saving ratio γ is given as follows:

$$\gamma = \frac{(C_A - C_S)E[N_{BL}]T_S + (C_A - C_L)E[N_{BL}]T_L}{E[T_R]} \tag{15}$$

where C_A, C_S and C_L are the energy consumption per slot in the awake state, the sleep state and the listen state, respectively.

We define the average response time δ of data frames as the time period in slots elapsed from the arrival of a data frame to the end of the transmission for this data frame, this measure can be used for evaluating the user Quality of Service (QoS). The expression for the average response time δ of data frames is obtained as follows:

$$\delta = P_L E[W_L] + P_A E[W_A] + E[S]. \tag{16}$$

Obviously, there is a trade-off between the energy saving ratio and the average response time of data frames when setting the time length for the sleep window. We develop an expected cost function $F(T_S)$ for a busy cycle R as follows:

$$F(T_S) = \frac{C_1}{E[N_{BL}]T_S} + C_2 E[N_{BL}]E[T_{BA}] + C_3\delta \tag{17}$$

where C_1 is the reward per slot when the MS is in the sleep state, C_2 is the cost per slot when the MS is in the awake state, and C_3 is the cost caused by the delay of data frames.

6 Numerical Results

According to [2], we set the parameters as follows: $E[S] = 2$ slots, $T_L = 4$ slots, $C_A = 30$ W, $C_L = 10$ W, and $C_S = 5$ W, where W is “watt". The numerical results in terms of the handover ratio β, the energy saving ratio γ and the average response time δ are presented in Figs. 1-3.

Figure 1 shows that how the handover ratio β changes with the system load ρ for different sleep window length T_S. It can be concluded that the handover ratio experiences a two stage. In the first stage, the handover ratio β will increase along with the increase of the system load ρ. During this stage, the larger the system load ρ is, the less possible the data frames will be finished transmitting in the listen state, then the more possible the system will enter into the awake state from the sleep state, so the handover ratio β will increase. In the second stage, the handover ratio β will decrease with the increase of the system load ρ. During this period, the larger the system load ρ is, the longer the MS will

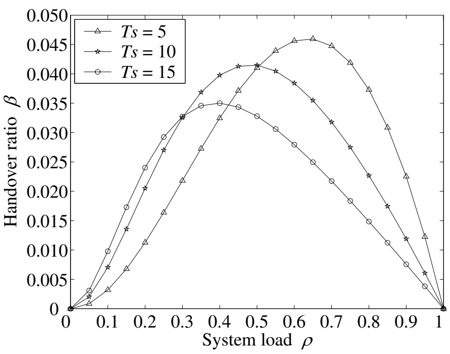

Figure 1: Handover ratio β vs. system load ρ.

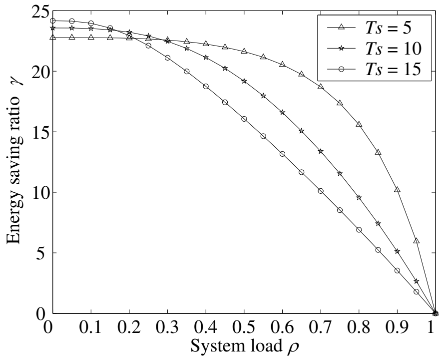

Figure 2: Energy saving ratio γ vs. system load ρ.

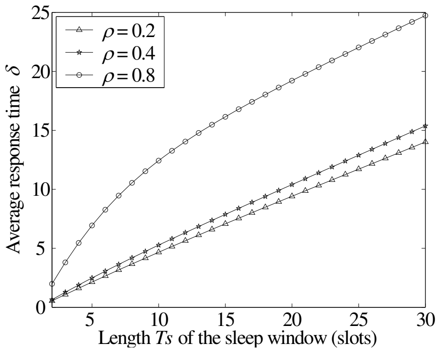

Figure 3: Average response time δ vs. sleep window length T_S (slots).

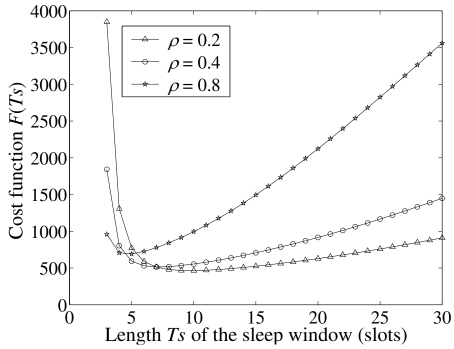

Figure 4: Cost function $F(T_S)$ vs. sleep window length T_S (slots).

stay in the awake state, so the less possible the system will switch to the sleep state from the awake state, and the lower the handover ratio β will be.

The influence of the system load ρ on the energy saving ratio γ with different sleep window length T_S is plotted in Fig. 2. It can be found that for the same length T_S of the sleep window, the energy saving ratio γ will decrease as the system load ρ increases. This is because the larger the system load ρ is, the less possible the data frames can be completely transmitted in the listen state, then the longer the MS will be in the awake state, so the less the energy saving ratio γ will be.

From Fig. 2, we can also find that for a less system load ρ, for example, $\rho < 0.2$, the larger the length T_S of the sleep window is, the larger the energy saving ratio γ is. The reason is that the larger the length T_S of the sleep window is, the longer the MS will stay in the sleep state, so the less the energy saving ratio γ will be. On the other hand, for a larger system load ρ, for example, $\rho > 0.2$, the larger the length T_S of the sleep window is, the less the energy saving ratio γ is. It is because that the less the length T_S of the sleep window is, the more possible the data frames arrived during the sleep window will be finished transmission in the listen state, the MS will more likely return to the sleep state from the listen state, so the larger the energy saving ratio γ will be.

Figure 3 examines the influence of the sleep window length T_S on the average response time δ of data frames. It can be observed that for the same system load ρ, the average response time δ will increase as the sleep window length T_S increases. This is because the longer the sleep window length T_S is, the longer the data frames arrived during the sleep state will wait in the sleep state, so the longer the average response time δ will be. For the same sleep window length T_S, the average response time δ will increase as the system load ρ increases. The reason is that the larger the system load ρ is, the much busier the system will be, and the longer the average response time δ will be.

Referencing [6], let C_1=3, $C_2 = 4$, and C_3=2 in Eq. (17), it can show how the cost function $F(T_S)$ changes with the sleep window length T_S for different system loads ρ in Fig. 4.

From Fig. 4, we can conclude that the cost function experiences two stage. In the first stage, the cost function $F(T_S)$ will decrease along with the increase of the sleep window length T_S. During this stage, the larger the sleep window length T_S is, the longer the MS will stay in the sleep state and the less the cost will be. In the second stage, the cost function $F(T_S)$ will increase with the increase of the sleep window length T_S. During this period, the larger the sleep window length T_S is, the larger the MS will stay in the sleep state, the longer the average response time δ is, so the larger the cost will be. Conclusively, there is a minimal cost for all the system loads when the sleep window length is set to an optimal value. For example, the optimal length of the sleep window is 4 slots when the system load $\rho = 0.2$, the optimal length of the sleep window is 7 slots when the system load $\rho = 0.4$, the optimal length of the sleep window is 9 slots when the system load $\rho = 0.8$.

7 Conclusions

Energy saving mechanism of the battery powered mobile stations is one of the most important issues for the application of the broadband wireless metropolitan area network (MAN). This paper proposed a novel method to analyze the system performance for the power saving class type II in IEEE 802.16e. A discrete-time queueing model with two kinds of vacation mechanisms was built to capture the working principle of the power saving class type II in this paper. With the performance analysis by using an embedded Markov chain and the boundary state variable theory, we gave the expressions about the performance measures for the power saving class type II in terms of the handover ratio, the energy saving ratio and the average response time. Finally, we presented numerical results to explain the nature of the dependency relationships between the performance measures and the system parameters, as well as developed a cost function to optimize the sleep window length under certain conditions. This paper provided a theoretical basis for the optimal setting of the system parameters in the power saving mechanisms, and has potential applications in solving other energy conserving related problems in wireless mobile networks.

Acknowledgements

This work was supported in part by National Science Foundation (No. 10671170, No. 60773100), Hebei Province National Science Foundation (No. F2009000475), China, and by GRANT-IN-AID FOR SCIENCE RESEARCH (No. 21500086) and the Hirao Taro Foundation of KUAAR, Japan.

References

[1] Y. Xiao, "Energy saving mechanism in the IEEE 802.16e wireless MAN," *IEEE Communications Letters*, vol. 9, pp. 595-597, 2005.

[2] B. Lee and H. Lee, "Cumulative-TIM method for the sleep mode in IEEE 802.16e wireless," *LNCS*, vol. 31, pp. 502-511, 2006.

[3] K. Han and S. Choi, "Performance analysis of sleep mode operation in IEEE 802.16e mobile broadband wireless access systems," in *Proc. of the IEEE VTC*, pp. 1141-1145, 2006.

[4] J. Lee and D. Cho, "An optimal power-saving class II for VoIP traffic and its performance evaluations in IEEE 802.16e," *Computer Communications*, vol. 31, pp. 3204-3208, 2008.

[5] N. Tian and Z. Zhang, *Queueing Models-Theory and Applications*, Springer, 2006.

[6] S. Jin and W. Yue, "Performance analysis for a system of connection oriented Internet service with a release delay," *IEICE Transactions on Communications*, vol. E90-B, pp. 3083-3094, 2007.