IEEE 802.16m Energy-Efficient Sleep Mode Operation Analysis with Mean Delay Restriction

Alexey Anisimov
H&NM Motorola
Software Organization, RUSSIA
alexey.anisimov@motorola.com

Sergey Andreev
State University of Aerospace
Instrumentation (SUAI), RUSSIA
serge.andreev@gmail.com

Andrey Turlikov
State University of Aerospace
Instrumentation (SUAI), RUSSIA
turlikov@vu.spb.ru

***Abstract*—IEEE 802.16e standard currently supports mobile subscriber stations, which substantiates the need for energy-efficient communication. In this paper we analyze a novel sleep mode mechanism to be considered for the future IEEE 802.16m standard. We study the overall message delay in the downlink channel and explicitly account for the sleep mode power consumption. Simple expressions are given based on M/D/1 queueing system with multiple vacations. The analytical approach is verified by simulation.**

***Index Terms*—IEEE 802.16m, sleep mode operation, average packet delay, power consumption, power saving.**

I. INTRODUCTION AND BACKGROUND

IEEE 802.16 group of standards defines a high speed wireless access system with support for various multimedia services. The standard contains specifications of Media Access Control layer (MAC) and of Physical layer (PHY). There are several possible realizations of the PHY layer. However the most popular one is Orthogonal Frequency Division Multiple Access (OFDMA). Originally IEEE 802.16 [1] has been designed for fixed Subscriber Stations (SS), but the most recent IEEE 802.16e [2] standard has support for mobile subscribers (MS). Therefore power saving is one of the constitutive issues and increasing MS energy efficiency is an important objective. To solve the issues the IEEE 802.16e standard defines the handover process and the sleep mode operation.

There are several proposals to make sleep mode operation more flexible and more effective [3], [4], [5]. These proposals can be included into the future IEEE 802.16m standard [6], which is currently being processed for standardization. The most part of research papers considers sleep mode defined by IEEE 802.16e standard. Zhang and Fujise [7] provide an analytical model to calculate the power consumption during the sleep mode. The model takes into account both types of traffic: downlink (from Base Station to MS) and uplink (from MS to BS). However the model estimates power consumption during a sleep period only (without awake periods) and message delay is not taken into account. In [8] the authors consider the average power consumption and the average delay under the sleep mode operation (including awake period). Moreover some information is provided about sleep mode parameters influence on the average delay and on the average power consumption.

The research closest to our work is summarized in [9]. The authors provide full analysis of the sleep mode operation for the efficient power saving mechanism and delay guaranteed services. They use the M/GI/1/K queueing system with multiple vacations. Moreover optimal parameters for sleep mode are presented. However this analysis considers the sleep mode operation of IEEE 802.16e standard. In [10] sleep mode operation proposed by Samsung for IEEE 802.16m is analyzed. In the paper only power consumption is considered without any delay analysis. One more opportunity to increase MS energy efficiency is scheduling mechanism improvement. This option is considered in [11].

In this paper we conduct the analysis of both delay and power consumption in case of sleep mode proposed for IEEE 802.16m. Our analysis applies M/D/1 queue with vacations. However IEEE 802.16 operation prohibits the direct use of already known results for M/D/1 queue with vacations. Thus, we consider a modified queueing model and derive the optimal parameters for the sleep mode.

The rest of the paper is structured as follows. In Section II we give a brief overview of sleep mode and its proposals. Section III provides the description of our model and notations. We consider the average delay and the average power consumption in Section IV and provide numerical results in Section V. We make a conclusion in Section VI.

II. IEEE 802.18m SLEEP MODE OPERATION

The sleep mode operation has three types of the Power Saving Classes (PSC). PSC is a group of connections that have common demand properties. PSC of type I is recommended for the connections of Best Effort (BE) and Non-Real-Time Variable-Rate (NRT-VR) service flows. PSC of type II is recommended for the connections of Unsolicited Grant Service (UGS) and Real-Time Variable-Rate (RT-VR) service flows. PSC of type III is recommended for multicast connections and for management flows. However, PSC of type III is optional [5].

In the sleep mode MS alternates sleep intervals and listening intervals (see Fig. 1). Before entering the sleep mode, the MS sends a request message ($MOB\_SLP-REQ$) to the BS for the permission to transit to sleep mode. After receiving the response message ($MOB\_SLP-RSP$) from the BS the MS switches the mode. To reduce the battery consumption the MS shuts down its radio-interface during sleep intervals and there is no data transmission between the BS and the MS. During listening interval the MS checks $MOB\_TRF-IND$ message.

9781-4244-3941-6/09/$25.00

It indicates the presence of any traffic addressed to the MS at the BS. If the MS received a negative $MOB\_TRF-IND$, i.e. no data for it, it goes back to the sleep state again.

In case of PSC of type II listening intervals and sleep intervals have constant duration. In case of PSC of type I the MS doubles sleep interval duration after each negative $MOB\_TRF-IND$ up to a maximum value. Therefore the following are relevant parameters in case of PSC of type I: initial sleep interval duration S_1, final sleep interval duration $S_{\max}$ and listening interval duration L. Denote sleep interval duration as S_i, where i is number of sleep cycle. In the case of PSC of type II $S_{i+1}=S_i$. In the case of PSC of type I:

$$S_i = \min\{2 \cdot S_{i-1}, S_{\max}\}. \quad (1)$$

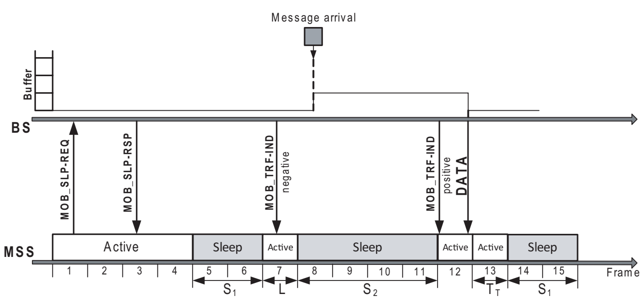

Fig. 1. Sleep mode operation example.

The MS continues the above procedure until it receives a positive $MOB\_TRF-IND$. After that the BS starts data transmission. According to [2] the MS sends $MOB\_SLP-REQ$ again to switch its state to sleep mode operation. It leads to unnecessary power expenses. To switch MS state back to sleep mode timer implementation is considered. When there are no packet arrivals during the following close-down time after the BS buffer empties, the MS enters the sleep mode state to decrease the power consumption. We denote timer duration by T_T.

III. SYSTEM MODEL

Our system model is based on [2], [3] and [4]. In this model we consider only PSC of type I and only downlink traffic is enabled. We use Poisson arrival process as a model of the incoming traffic. The BS has an infinite buffer capacity to store the waiting messages. We apply a discrete-time model, in which the time is slotted. Each slot is equal to the transmission time of a data message. However, all the time durations are measured between the relevant events in real-value. This slight modification of the discrete-time model ensures that the analytical results fit those of the corresponding continuous-time model.

The considered sleep mode mechanism uses a close-down time interval to switch to the sleep state. There are no extra transmissions of MOB_SLP messages. When there are no data transmissions during the close-down time, the MS enters the sleep state with the state duration equal to S_1.

When packet arrival occurs during sleep interval it waits for the listening interval. After that it is transmitted. Note if a packet arrives at the BS during frame number i, it could be transmitted not earlier than in the next frame, i.e. in frame number $i+1$. In particular, if a packet arrives during frame number 7 (see Fig. 1), it will be transmitted in frame number 12.

In our model we apply First Come First Serve (FCFS) discipline. We consider only one MS in the system. So it means that there are enough resources to transmit $MOB\_TRF-IND$ and data messages at listening intervals in downlink channel. Therefore data messages are transmitted frame by frame without interruptions until the buffer empties.

IV. ANALYTICAL APPROACH

Our analysis is based on fundamental works such as [12], [13]. In addition we use a mechanism which is considered in [14]. Data transmission process from the BS to the MS can be considered as a regenerative process. The regenerative process has the property of 'starting afresh probabilistically' from time to time. We divide the time axis into intervals which are called regeneration cycles. The cycle begins at the same time when the close-down timer starts, if there are no any data transmissions during T_T time period. The interval ends when the BS buffer becomes empty (see Fig. 2). There is at least one data transmission during a regeneration cycle. To estimate the mean message delay and power consumption it is enough to calculate them for one regeneration cycle only.

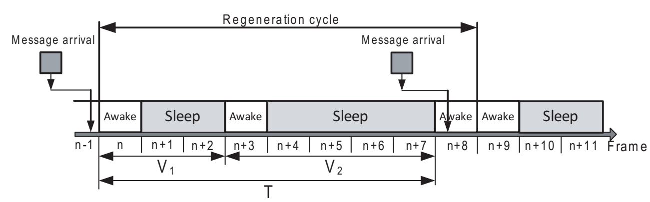

Fig. 2. An example of the regeneration cycle.

A. Transmission delay

Our system is a M/D/1 system with vacations. In [13] an analysis to estimate the mean delay is provided. The analysis is applicable in case of an independent identically distributed vacation lengths. The authors note that the approach is applicable in case of dependent identically distributed vacation lengths as well. However in our case vacation lengths have different distributions. We extended the derivation from [13] for our model and obtained the following expression:

$$E[W] = \frac{\lambda \cdot \bar{X}^2}{2 \cdot (1-\rho)} + \frac{\bar{T}^{(2)}}{2 \cdot \bar{T}} + W_t, \quad (2)$$

where λ is overall arrival rate, $\bar{X}$ is the service time, ρ is utilization factor, $\bar{T}$ is the estimation of time interval from the start of the regeneration cycle up to the first data transmission and $\bar{T}^{(2)}$ is the second moment of it. W_t is a transmission delay of a message.

Expression (2) is formally close to Pollaczek-Khinchine formula. However the second term has the different meaning. In the original case of the expression for M/D/1 queue with vacations this term is calculated via durations of vacation intervals, which we denote by S_i. To estimate the overall message delay in case of IEEE 802.16 model we use time interval from the beginning of the regeneration cycle up to the first frame in which any data transmission occurs, i.e. T.

Here we provide an approach to calculate $\bar{T}^{(2)}$ and $\bar{T}$. In our case the vacation length equals to the following:

$$\begin{aligned} V_1 &= T_f + S_1 \\ V_2 &= L + S_2 \\ &\dots \\ V_i &= L + S_i. \end{aligned} \quad (3)$$

We denote possible regeneration cycle length by t_i.

$$\begin{aligned} t_0 &= 0 \\ t_1 &= V_1 \\ t_2 &= t_1 + V_2 \\ &\dots \\ t_i &= t_{i-1} + V_i. \end{aligned} \quad (4)$$

Expressions to calculate $\bar{T}$ and $\bar{T}^{(2)}$ are below.

$$\bar{T} = \sum_{i=1}^{\infty} \Pr\{T = t_i\} \cdot t_i, \text{ and}$$

$$\bar{T}^{(2)} = \sum_{i=1}^{\infty} \Pr\{T = t_i\} \cdot t_i^2, \quad (5)$$

where $\Pr\{T = t_i\} = \Pi_0^{t_{i-1}} \cdot (1 - \Pi_0^{V_i})$ and $\Pi_{t_i}^0$ is probability that there are no message arrivals during t_i. In the case of Poisson arrivals it equals to:

$$\Pi_0^{t_i} = e^{-t_i \cdot \lambda}. \quad (6)$$

B. Energy efficiency analysis

We define energy efficiency as the following:

$$U = \frac{E[\text{transmitted data}]}{E[\text{power consumption}]}. \quad (7)$$

To obtain the U value it is enough to estimate the quantity of transmitted data bits per regeneration cycle and power consumption per regeneration cycle. Therefore, we define the quantity of transmitted data bits and power consumption as follows:

$$E[\text{transmitted data}] = b \cdot \sum_{i=1}^{\infty} f_i \cdot i, \text{ and}$$

$$E[\text{power consumption}] = P_A \cdot T_A + P_S \cdot T_S, \quad (8)$$

where i is number of transmitted data messages per regeneration cycle, f_i is probability that i messages are transmitted per regeneration cycle, b is the size of the message. P_A and P_S are power consumption in active and sleep state per unit of time respectively. T_A and T_S stand for time spent in active state and sleep state respectively.

In [12] a formula for f_i calculation is provided. However the formula is applicable only when there is one message in the buffer at the beginning of an awake period. In our case there could be several messages. Therefore we generalize the formula for the case of several messages.

Denote number of messages which are stored in the buffer at the beginning of the awake period by X

$$f_i = \sum_{x=1}^{i} \Pr\{X = x\} \cdot A_{i-x}^{x}, \quad (9)$$

where $\Pr\{X = x\}$ is probability that there are x messages in the buffer at the beginning of the awake period. A_{i-x}^{x} is probability that $i-x$ messages will arrive during transmission time of x messages.

$$\Pr\{X = x\} = \sum_{i=0}^{\infty} \Pi_0^{t_i} \cdot \Pi_x^{V_{i+1}}, \quad (10)$$

where $\Pi_0^{t_i}$ is probability that there are no message arrivals during t_i time interval. $\Pi_x^{V_{i+1}}$ is probability that there are x message arrivals during V_{i+1} time interval.

$$A_c^b = \sum_{i=1}^{c} \Pr\{X = i\} \cdot A_{c-i}^{i}, \text{ and}$$

$$A_0^i = e^{-i \cdot \lambda}. \quad (11)$$

Here we consider active and sleep time intervals.

$$T_S = \sum_{i=1}^{\infty} S_i \cdot \Pi_0^{t_{i-1}} \cdot (1 - \Pi_0^{S_i + L}), \text{ and}$$

$$T_A = f_i + \sum_{i=1}^{\infty} i \cdot \Pi_0^{t_{i-1}} \cdot (1 - \Pi_0^{S_i + L}). \quad (12)$$

V. NUMERICAL RESULTS

In order to validate the considered analytical model a simulation program for IEEE 802.16m MAC was developed. To obtain numerical results we use recommendations which are provided by [2], [15]. We set the simulation parameters of IEEE 802.16m MAC and PHY as the following:

Basic simulation parameters:

Parameter	**Value**
PHY layer	OFDMA
Frame duration, T_f	5 ms
Listening interval, L	1 $frame$
Time-out interval, T_T	1 $frame$
Power consumption in active state, P_A	750 mW
Power consumption in sleep state, P_S	50 mW
Packet size, b	1536 $bits$

Our analysis is applicable for M/D/1 system with vacations. However, we use HTTP arrival model to show the difference with the Poisson traffic. The HTTP traffic model is based on [16]. According to the description, the traffic source has two

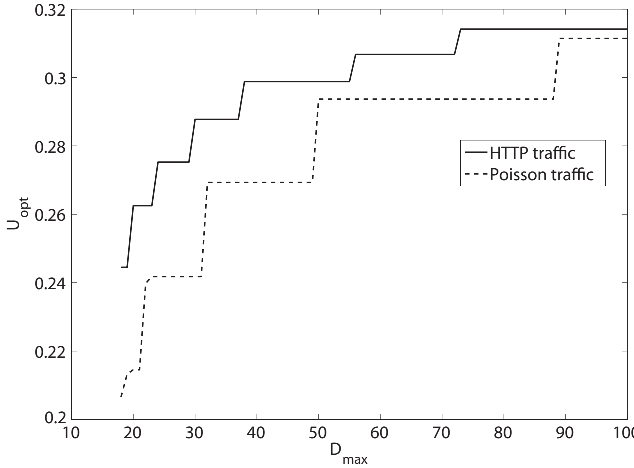

Fig. 3. Optimal energy efficiency vs. maximum tolerable delay.

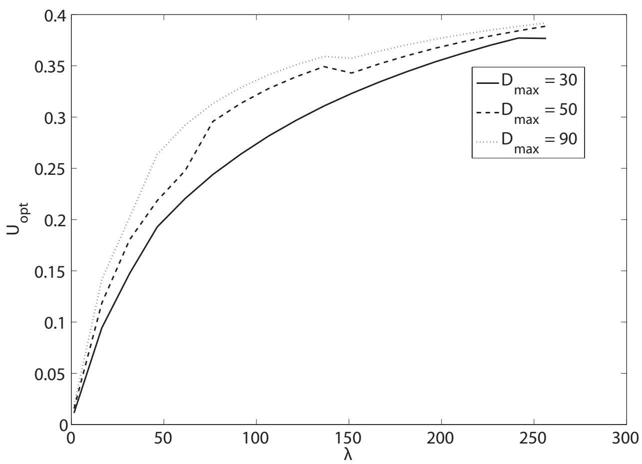

Fig. 4. Optimal energy efficiency vs. arrival rate.

states: ON and OFF. In the ON state the source generates new messages, while in the OFF state it is silent. This model is also known as Interrupted Poisson Process. We may, therefore, consider two arrival rates: arrival rate in the ON state and the average arrival rate. In Fig. 3 we plot the Poisson traffic curve with the rate equal to the average arrival rate for the HTTP model.

Fig. 3 demonstrates the dependence of the optimal energy efficient value on the maximum tolerable delay. We see that the optimal energy efficiency generally increases with the growth of the D_{max}. We also emphasize that our model allows for reaching approximately the same energy-efficient level. Each step corresponds to another set of sleep mode parameters (initial and final sleep interval durations). The difference between the steps corresponds to the delay discretization gap, which is the delay value we may reach subject to the maximum energy efficiency. We see that the mean discretization gap is lower in case of HTTP traffic, than in case of Poisson traffic.

Fig. 4 compares optimal energy efficiencies on the arrival rate for different D_{max} values. As follows from the energy efficiency definition, the optimum is the higher, the greater is the arrival rate. At the same time the higher the tolerable delay is, the greater energy efficiency we may attain. Fig. 4 shows that the main difference in the optimal energy efficient levels is noticed for moderate arrival rates (40-150 Kbps). When the arrival rate is low or near-critical, the optimal energy efficiency is almost independent of the mean delay constraint.

VI. CONCLUSION

In this paper we developed an analytical discrete-time model to evaluate the average overall delay and the power consumption of the future IEEE 802.16m wireless network in case of the sleep mode operation and provided simple expressions for these values. The model allows for Poisson input traffic and FCFS service discipline.

It is clear that longer sleep windows lead to higher energy efficiency. However, at the same time, they increase the overall message delay. When the overall message delay is limited, with our approach it is possible to set optimal sleep mode parameters, which maximize the energy efficiency.

Moreover we provided the comparison of Poisson arrival model with HTTP traffic. This comparison shows that our analysis can be used as a lower bound estimation for energy efficiency in the case of HTTP traffic. Also we showed a dependence of achievable energy efficiency on maximum tolerable delay.

REFERENCES

[1] *IEEE 802.16-2004, Part 16: Air Interface for Fixed Broadband Wireless Access Systems, October 2004.*
[2] *IEEE 802.16e-2006, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, February 2006.*
[3] *Sleep Mode Operation for IEEE 802.16m, C802.16m-08/721r1, July 2008.*
[4] *Keep-awake Mechanism for IEEE 802.16m Sleep Mode, C802.16m-08/718, July 2008.*
[5] *Sleep Mode for IEEE 802.16m System, C802.16m-08/688r1, July 2008.*
[6] *The Draft IEEE 802.16m System Description Document, IEEE 802.16m-08/003r5, October 2008.*
[7] Y. Zhang and M. Fujise, "Energy management in the ieee 802.16e mac," *IEEE Communications Letters*, vol. 10, no. 4, pp. 311–313, 2006.
[8] K. Han and S. Choi, "Performance analysis of sleep mode operation in ieee 802.16e mobile broadband wireless access systems," *IEEE VTC*, vol. 3, pp. 1141–1145, 2006.
[9] Y. Park and G. U. Hwang, "An efficient power saving mechanism for delay-guaranteed services in ieee 802.16e," *IEEE VTC*, 2007.
[10] S. Baek, J. J. Son, and B. D. Choi, "Performance analysis of sleep mode operation for ieee 802.16m advanced wman," *MACOM*, 2009.
[11] J. Y. Kim, T. Kwon, and D.-H. Cho, "Resource allocation scheme for minimizing power consumption in ofdm multicast systems," *IEEE Communications Letters*, vol. 11, no. 6, pp. 486–488, 2007.
[12] L. Kleinrock, *Queueing Systems, volume 1: Theory*, New York, 1975.
[13] D. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, 1992.
[14] M. A. Crane and A. J. Lemoine, *An Introduction to the Regenerative Method for Simulation Analysis*, Berlin - Heidelberg - New York, 1977.
[15] *WiMAX System Evaluation Methodology, Version 2.1, WiMAX Forum, July 2008.*
[16] *Traffic Model for 802.16 TG3 MAC-PHY Simulations, IEEE 802.16.3c-01/30r1, March 2001.*