Modeling and Analysis of Watchful Sleep Mode With Different Sleep Period Variation Patterns in PON Power Management

Min Zhu, Xiaobo Zeng, Yang Lin, and Xiaohan Sun

***Abstract*—Reducing power consumption in an access network has become an increasingly imperative design goal, due to the fact that information and telecommunication technologies contribute to an increasingly large proportion of greenhouse gas emissions. The passive optical network (PON) is considered as the most attractive and promising technology to provide low-cost services to end users in a power-efficient way due to the passive nature of remote nodes. However, it is necessary to further reduce the energy consumption of PONs, with the wide deployment of PONs and the rapid growth of data transfer rates in PONs. In this paper, we focus on the watchful sleep mode with different sleep period variation patterns for multiple optical network units (ONUs). This is because, in the traditional operational mode, ONUs have to continually listen to and inspect traffic from the OLT and hence always remain active even when there is no/light traffic, which contributes to the majority of energy wastage to the PON. We first modeled the watchful sleep mode based on the Markov chain model in order to analyze the effect of each key parameter on system performances in terms of the energy-saving efficiency and data packet delay. Due to the fact that the *sleep* state is the key to power-saving, we designed four different sleep period variation patterns (e.g., *constant*, *linear*_1, *linear*_2, and *exponential* patterns) to study the impact of these different patterns on the integrated performance in terms of the normalized *cost* value. Through extensive simulations, we found that, in the watchful sleep mode, the effect of the number (*n*) of *(sleep, listen)* state pairs would be insignificant, and the effects of other parameters on the performances are analyzed comprehensively. The minimum normalized *cost* value can be obtained under the four different sleep period variation patterns, which represent the optimal trade-off between the above two conflicting performances indexes.**

***Index Terms*—Energy-saving efficiency; Markov chain model; Passive optical networks; Sleep period variation pattern; Watchful sleep mode.**

I. Introduction

In the grim reality of high energy costs, academic and industry efforts have been focused on low energy consumption technologies [1,2]. The Internet and associated information and telecommunication technologies (ICTs) are playing an important role in our daily lives. However, ICTs are accountable for about 2%–4% of the worldwide greenhouse gas emissions [3–5]. It is predictable that the situation would worsen with a wider deployment of ICTs. In the last mile, various access technologies were proposed such as WiMAX, FTTH, point-to-point optical access networks, and so on [6–8]. The passive optical network (PON) stands out among these technologies and is considered the most attractive and promising technology to provide low-cost services to end users in a power-efficient way due to the passive nature of the remote nodes [9,10].

The PON is a point-to-multipoint optical network, as shown in Fig. 1. The optical line terminal (OLT) resides at the central office (CO). It is connected to multiple optical network units (ONUs) through one or multiple passive optical splitters [11]. Generally, the PON provides a basic optical fiber infrastructure without having an external power supply outside the CO. Thus, it considerably reduces power consumption and the operational cost. In addition, the advantages of PONs include high bandwidth, long distance coverage, low cost, and transparent upgrade capability [3,12].

Nevertheless, owing to the rapid growth of data transfer rates of PON systems, reducing the energy consumption of PON as much as possible is essential and imperative for sustainable development. It is due to the fact that the larger data transfer rate needs the support of more energy to drive the optical and electrical components of various devices [9].

Note that the ONUs located at the user-side cause the majority of energy wastage. That is due to the characteristic of PON that the ONUs have to continually listen and inspect traffic from the OLT. Hence, the ONUs always remain active even when there is no/light traffic. Therefore, further reducing the energy consumption of ONUs is the key to achieving the power-saving goal in PON [3,13].

Until now, the doze mode and the cyclic sleep mode were standardized by ITU-T G.987.3 for the ONU power management [14]. They are operated by turning OFF/ON all or part of the transceivers (i.e., transmitter and receiver) in the ONUs. Thus, the ONU realizes the state switching between the FULL power state (e.g., the *active held* and the

Manuscript received February 10, 2017; revised July 12, 2017; accepted July 12, 2017; published September 1, 2017 (Doc. ID 286386).

M. Zhu (e-mail: minzhu@seu.edu.cn) and X. Sun are with the National Research Center for Optical Sensing/Communications Integrated Networking, School of Electronic Science and Engineering, Southeast University, Nanjing 210096, China.

X. Zeng and Y. Lin are with the School of Microelectronics, Southeast University, Nanjing 210096, China.

https://doi.org/10.1364/JOCN.9.000803

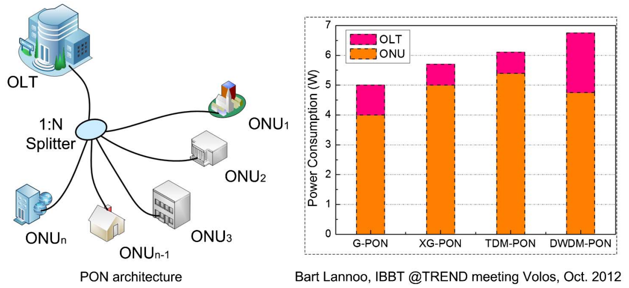

Fig. 1. PON architecture and power consumption for the OLT and ONU.

active free states) and LOW power state. Note that the LOW power state represents the *listen* state and the *sleep* state for the doze mode and the cyclic sleep mode, respectively [15,16]. In the above two modes, there still exists the transition states (i.e., *sleep aware* and *doze aware*, respectively) to achieve the synchronization between the OLT and ONUs, through the exchange of signaling information.

Bang *et al.* optimized the energy-saving efficiency of the cyclic sleep mode, without impairing the data packet delay by using the state probability analysis model [17]. Zhang *et al.* explored the energy efficiency of ONUs for the sleep control scheme [9]. Khotimsky *et al.* unified the standardized doze mode and cyclic sleep mode into a new mode, which was named as the watchful sleep mode [1,16]. The new mode simplifies the operation of the energy-saving mode in ONUs. It can emulate the doze mode and the cyclic sleep mode as special cases.

For the watchful sleep mode, some preliminary results are given out in our previous works [18,15]. We found that the sleep period is a significant parameter, which greatly affects the PON system performances: energy-saving efficiency (ESE) and data packet delay (DPD). Furthermore, there is a trade-off between these two conflicting indices for the sleep period [18]. The larger the sleep period is, the more time the ONU spends in the *sleep* state. Thus, it is favorable that the energy-saving efficiency of the PON becomes higher. However, the data packet delay would turn larger.

In this paper, we further modeled the watchful sleep mode for ONUs based on the Markov chains with the complete mathematical analyses and expressions. Then we performed extensive numerical simulations to study the effects of all key parameters on the system performances in terms of the ESE and the DPD [15,18]. Moreover, we designed several different sleep period variation patterns to investigate the impacts of the extended sleep periods on the system-integrated performance (i.e., normalized *cost* value).

The rest of the paper is organized as follows. The Markov chain model was detailed for the watchful sleep mode in Section II. We theoretically analyze the performances of the ESE and the DPD in Section III. Section IV analyzes the impact of some relevant parameters on the system performances, based on the obtained numerical results. Section V concludes the paper.

II. Mathematical Model Based on the Markov Chain for the Watchful Sleep Mode

In this section, we will first describe the operation principle of the watchful sleep mode in Subsection II.A. Then the state transition model for the state machine of the ONU will be established based on the Markov chain in Subsection II.B. As a result, the steady-state probabilities for all the possible states are derived in Subsection II.C. At last, four different sleep period variation patterns for the watchful sleep mode are proposed in Subsection II.D.

A. Operation Principles of Watchful Sleep Mode for Power Management in ONUs

The watchful sleep mode at the ONU is operated based on the five power management states: *active held*, *active free*, *aware*, *listen*, and *sleep* states, as shown in Fig. 2. The first two states constitute the active phase, while the later three states comprise the power-saving phase of the ONU. When the ONU stays in the active phase, its transmitter (TX) and receiver (RX) are fully responsive and forwarding the upstream (US) or downstream (DS) traffic. It should be noted that the ONU cannot directly enter the power-saving phase from the *active held* state. And the ONU, which stays in the *active free* state, can freely enter the power-saving phase only if there is no traffic in the whole duration time of the *active free* state. Alternatively, the ONU can return to the *active held* state from the *active free* state if it received the wakeup indication signal, as indicated in Fig. 2. During the power-saving phase, the ONU operates in an outer circulation between the full power *aware* state (TX and RX are ON) and the low power *watch* state. Further, in the *watch* state, the ONU switches in an inter-circulation between the *listen* state and *sleep* state (see dashed block), as shown in Fig. 2. When there exists arrival data traffic during the *aware* or *listen* state, a wake-up indication occurs, and the ONU would directly get out of the power-saving phase and enter into the *active held* state. If the data traffic arrived during the sleep period, which cannot be interrupted, these

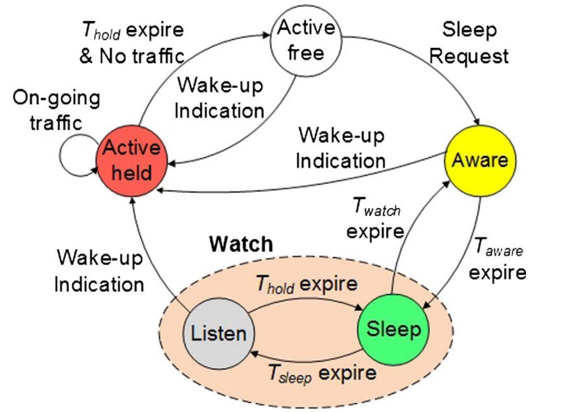

Fig. 2. State machine of ONU for the *watchful* sleep mode.

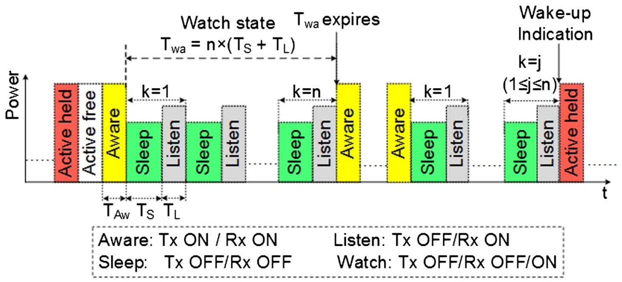

Fig. 3. Power consumption of the ONU.

data are cached at the buffer queue of the ONU. At the end of the sleep period, the ONU can transit to the *active held* state then forward these cached data.

The duration time of the *listen* state and *sleep* state are denoted as T_L and T_S, respectively. It is assumed that the *watch* state contains n pairs of the (*sleep*, *listen*) state pairs; thus, $T_{\mathrm{wa}} = n \times (T_S + T_L)$, where T_{wa} is the duration time of the *watch* state, as shown in Fig. 3. In the *watch* state, the TX of ONU is always OFF, while the RX is periodically ON for a short time to detect for the wakeup indication and handle the arrived packets, thus creating two possible power levels: I) the TX and RX are OFF (e.g., the *sleep* state), and II) the TX is OFF, but the RX remains ON (e.g., the *listen* state). When the ONU has not received any traffic packet until T_{wa} expires, the ONU will return to the *aware* state to handle the necessary synchronization with the OLT.

B. State Transitions Based on the Markov Chain

The Markov chain is a random process that undergoes the transition from one to another state within the state space. The random process must satisfy the Markov property that the probability distribution of the next state depends only on the current state and not on the sequence of events that preceded it [19].

In this paper, we employed the Markov chain to analyze the ONU performances of energy-saving efficiency and the data packet delay from a theoretical perspective, and attempt to obtain an optimal trade-off between the above two conflicting performance indexes. This paper just considers the downstream traffic scenario because the upstream traffic is generally buffered in the ONU and transmitted to the OLT within the allocated time window in a time division multiplexing way. However, the upstream traffic scheduling is beyond the scope of this paper. The Poisson traffic arrival process was adopted for the sake of simplifying the analysis. The packet arrival process was assumed to be independent for the distribution of packet size. The average traffic arrival rate was denoted as λ packets/s, which refers to the number of data packets arriving within a unit time. The service time of each packet follows the exponential distribution with the average value of $1/\mu$. Note that λ/μ is defined as the traffic load, which reflects the free/busy status of the ONUs. When the ONU is free, the ONU can enter into the power-saving phase. In the Markov chain model for the proposed state machine of the ONU, the ONU has five possible state sets (as follows), and each state is divided into several scheduling cycles (i.e., substates). One substate lasts for one scheduling cycle (i.e., $T = 2$ ms). In this paper, we do not consider the transit time between any two states because the transit time (i.e., tens of nanoseconds) can be negligible compared with the duration of a scheduling cycle (i.e., $T = 2$ ms).

- *Active held* state Ah_i ($i = 1, 2, 3, \ldots$) represents the state sets that the ONU stays in the *active held* state. There exits i queued traffic packets by the end of one scheduling cycle, and the upper bound of the value of i is theoretically infinite. The ONU in the Ah_i substate can transfer into any other FULL power substates (i.e., Af, Ah_j ($j \neq i$)) except those substates of the *Aware* state, when the number i of the queued traffic in the buffer was changed.
- *Active free* state Af is a transition state with one scheduling cycle between the *active held* and the *aware* state. Once a wakeup indication occurs during its scheduling cycle, the ONU can switch into the *active held* state. If there is not any traffic arrival during the scheduling cycle, the ONU can transfer into the *aware* state.
- *Aware* state Aw_j ($j = 1, 2, \ldots, x$) is a FULL power state in the low power phase, where the ONU can handle some necessary synchronizations with the OLT. Generally, the ONU cannot return to the Af state but can directly transfer into the Ah_i state with the trigger of the wakeup indication.

In addition, the *watch* state is composed of the *listen* state and the *sleep* state, which are described in detail as follows:

- *Listen* state $L_{k,l}$ ($k = 1, 2, \ldots, n; l = 1, 2, \ldots, y$) is the lth substate of the *listen* state in the kth (*sleep*, *listen*) state pair. If an ONU stays in the $L_{k,l}$ substate, this means that the ONU has not received any data packet for $(k - 1) \times y + l$ scheduling cycles. Once the existing data packets arrive in the ONU during the scheduling cycle of the $L_{k,l}$ substate, a wake-up indication occurs, and the ONU can transfer into the Ah_i ($i = 1, 2, 3, \ldots$) substate. Otherwise, the ONU gets into the *sleep* state if it has still not received any traffic packets during the remaining $(y - l)$ scheduling cycles and $k \neq n$. In addition, the ONU would also transfer into the *aware* state when T_{wa} was expired (i.e., $k = n$ and $l = y$), as shown in Fig. 3.
- *Sleep* state S_k ($k = 1, 2, \ldots, n$) refers to those substates where the ONU stays asleep. It is noted that the duration time T_S of each S_k is z cycles (i.e., $T_S = z * 2$ ms), which cannot be interrupted until T_S is expired. At the end of the *sleep* state, the ONU can transfer into the Ah_i ($i = 1, 2, 3, \ldots$) state if it received a wake-up indication from the OLT to process those traffic packets that arrived in the period of the *sleep* state. Otherwise, the ONU can transfer into the *listen* state, if there is not any packet arrival during the *sleep* state.

Figure 4 illustrates the state transition among all the above states. Based on the assumptions of Poisson arrival

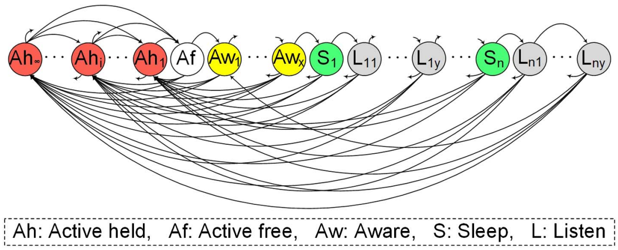

Fig. 4. State transition of the ONU in the Markov chain.

traffic, for a given ONU, the probabilities of the arrival of b DS packets at the ONU and the departure of b US packets from the ONU within g scheduling cycles can be represented as $p^a(b,g)$ or $p^d(b,g)$ as follows, respectively:

$$p^a(b,g) = e^{-\lambda gT}(\lambda gT)^b/b!, \tag{1}$$

$$p^d(b,g) = e^{-\mu gT}(\mu gT)^b/b!. \tag{2}$$

Thus, according to the Markov chain model, we could detail the following cases with regard to the state transitions.

1) If there is no packet arrival during one scheduling cycle, the state transition from Af to Aw_1, from Aw_j to Aw_{j+1}, from Aw_x to S_1, from $L_{k,l}$ to $L_{k,l+1}$, and from $L_{k,y}$ to S_{k+1} would have happened, and the probability of each case, as noted above, can be written as $p^a(0,1)$. If there is no packet arrival during z *sleep* cycles, the state transition occurs from the S_k to the $L_{k,1}$, and the probability is $p^a(0,z)$.
2) If the number of arrival packets is greater than the number of departures by k during one cycle, the state transition from $L_{k,l}$, Af, Aw_j to Ah_i, from Ah_i to Ah_{i+k} would have happened, and the probability is $\Sigma_i p^a(k+i,1)p^d(i,1)$. If the above case happens in the z cycles of the *sleep* state, the ONU state can transit from S_k to Ah_k, and the probability equals $\Sigma_i p^a(k+i,z)p^d(i,z)$.
3) If the number of arrival packets is smaller than the number of departures by k during one cycle, the state transition probability from Ah_i to Ah_{i-k} would have happened, and the probability is $\Sigma_i p^a(i,1)p^d(k+i,1)$.

C. Steady-State Probability

According to the principle of the balance equations in the Markov chains, the arrival frequency of a certain state equals the departure frequency of the state [20]. Thus, the steady-state probability of all the substates $P(\mathrm{Ah}_i)$ $(i=1,2,3,\ldots)$, $P(\mathrm{Af})$, $P(\mathrm{Aw}_j)$ $(j=1,2,\ldots,x)$, $P(L_{k,l})$ $(k=1,2,\ldots,n; l=1,2,\ldots,y)$, and $P(S_k)$ $(k=1,2,\ldots,n)$ must satisfy the constraints as follows.

The steady-state probability for all substates in (*sleep*, *listen*) state pairs are described as follows, except for the S_1 and $L_{n,y}$ state:

$$P(L_{k,y})\left[\sum_{i=1}^{\infty}\Pr\{L_{k,y}\to \mathrm{Ah}_i\} + \Pr\{L_{k,y}\to S_{k+1}\}\right] = P(L_{k,(y-1)})[\Pr\{L_{k,(y-1)}\to L_{k,y}\}], \tag{3}$$

$$P(L_{k,l})\left[\sum_{i=1}^{\infty}\Pr\{L_{k,l}\to \mathrm{Ah}_i\} + \Pr\{L_{k,l}\to L_{k,(l+1)}\}\right] = P(L_{k,(l-1)})[\Pr\{L_{k,(l-1)}\to L_{k,l}\}], \tag{4}$$

$$P(L_{k,1})\left[\sum_{i=1}^{\infty}\Pr\{L_{k,1}\to \mathrm{Ah}_i\} + \Pr\{L_{k,1}\to L_{k,2}\}\right] = P(S_k)[\Pr\{S_k\to L_{k,1}\}], \tag{5}$$

$$P(S_k)\left[\sum_{i=1}^{\infty}\Pr\{S_k\to \mathrm{Ah}_i\} + \Pr\{S_k\to L_{k,1}\}\right] = P(L_{(k-1),y})[\Pr\{L_{(k-1),y}\to S_k\}]. \tag{6}$$

The steady-state probability of S_1 equals

$$P(S_1)\left[\sum_{i=1}^{\infty}\Pr\{S_1\to \mathrm{Ah}_i\} + \Pr\{S_1\to L_{1,1}\}\right] = P(\mathrm{Aw}_x)[\Pr\{\mathrm{Aw}_x\to S_1\}]. \tag{7}$$

The steady-state probability of $L_{n,y}$ is defined as follows:

$$P(L_{n,y})\left[\sum_{i=1}^{\infty}\Pr\{L_{n,y}\to \mathrm{Ah}_i\} + \Pr\{L_{n,y}\to \mathrm{Aw}_1\}\right] = P(L_{n,(y-1)})[\Pr\{L_{n,(y-1)}\to L_{n,y}\}]. \tag{8}$$

The following equations are the steady-state probabilities of the *aware* substates:

$$P(\mathrm{Aw}_j)\left[\sum_{i=1}^{\infty}\Pr\{\mathrm{Aw}_j\to \mathrm{Ah}_i\} + \Pr\{\mathrm{Aw}_j\to \mathrm{Aw}_{(j+1)}\}\right] = P(\mathrm{Aw}_{(j-1)})[\Pr\{\mathrm{Aw}_{(j-1)}\to \mathrm{Aw}_j\}], \tag{9}$$

$$P(\mathrm{Aw}_x)\left[\sum_{i=1}^{\infty}\Pr\{\mathrm{Aw}_x\to \mathrm{Ah}_i\} + \Pr\{\mathrm{Aw}_x\to S_1\}\right] = P(\mathrm{Aw}_{(x-1)})[\Pr\{\mathrm{Aw}_{(x-1)}\to \mathrm{Aw}_x\}], \tag{10}$$

$$P(\mathrm{Aw}_1)\left[\sum_{i=1}^{\infty}\Pr\{\mathrm{Aw}_1\to \mathrm{Ah}_i\} + \Pr\{\mathrm{Aw}_1\to \mathrm{Aw}_2\}\right] = P(\mathrm{Af})[\Pr\{\mathrm{Af}\to \mathrm{Aw}_1\}] + P(L_{n,y})[\Pr\{L_{n,y}\to \mathrm{Aw}_1\}]. \tag{11}$$

In addition, the steady-state probability of Af equals

$$P(\mathrm{Af})\left[\sum_{i=1}^{\infty}\Pr\{\mathrm{Af}\to \mathrm{Ah}_i\} + \Pr\{\mathrm{Af}\to \mathrm{Aw}_1\}\right] = \sum_{i=1}^{\infty}P(\mathrm{Ah}_i)[\Pr\{\mathrm{Ah}_i\to \mathrm{Af}\}]. \tag{12}$$

The steady-state probabilities of the *active held* substates are expressed as follows:

$$\begin{aligned}P(\mathrm{Ah}_i)&\left[\sum_{I\neq i}^{\infty}\Pr\{\mathrm{Ah}_i\to\mathrm{Ah}_I\}+\Pr\{\mathrm{Ah}_i\to\mathrm{Af}\}\right]\\&=\sum_{I\neq i}^{\infty}P(\mathrm{Ah}_I)[\Pr\{\mathrm{Ah}_I\to\mathrm{Ah}_i\}]+\sum_{j=1}^{x}P(\mathrm{Aw}_j)[\Pr\{\mathrm{Aw}_j\to\mathrm{Ah}_i\}]\\&\quad+P(\mathrm{Af})[\Pr\{\mathrm{Af}\to\mathrm{Ah}_i\}]+\sum_{k=1}^{n}\left[\sum_{l=1}^{y}P(L_{k,l})\cdot[\Pr\{L_{k,l}\to\mathrm{Ah}_i\}]\right]\\&\quad+\sum_{k=1}^{n}P(S_k)[\Pr\{S_k\to\mathrm{Ah}_1\}].\end{aligned}\tag{13}$$

Furthermore, the sum of probabilities of all the substates equals 1.

$$\begin{aligned}\{P(\mathrm{Ah}_i)\}_{i=1}^{+\infty}&+P(\mathrm{Af})+\{P(\mathrm{Aw}_j)\}_{j=1}^{x}\\&+\{\{P(L_{k,l})\}_{l=1}^{y}\}_{k=1}^{n}+\{P(S_k)\}_{k=1}^{n}=1.\end{aligned}\tag{14}$$

So far, all of the state probabilities can be calculated from Eqs. (1)–(14).

D. Sleep Period Variation Patterns

It is noted that the duration time and the variation pattern of the *sleep* state are critical for the ONU performances in the watchful sleep mode. Hence, we design four different sleep period variation patterns and study their impacts on performances in terms of the energy-saving efficiency and the data packet delay. We try to obtain an optimal trade-off between the two conflicting performance indexes. The duration T_S of the *sleep* state can be expressed as $T_S = z \times 2$ ms. Four different sleep period variation patterns can be distinguished by the relationship of the following two parameters: the number z of scheduling cycles in the *sleep* state and the number n of the state pair (*sleep*, *listen*). Specifically,

1) ***Constant*** pattern: z is taken to be a fixed value, which is independent of n ($n = 1, 2, \ldots$).
2) ***Linear_1*** pattern: z linearly increases as n increases and satisfies the equation of $z = n$ ($n = 1, 2, \ldots$).
3) ***Linear_2*** pattern: z linearly increases as n increases and follows the relationship of $z = 2n$ ($n = 1, 2, \ldots$).
4) ***Exponential*** pattern: z increases exponentially as n increases and follows the equation $z = 2^{n-1}$ ($n = 1, 2, \ldots$).

III. Performance Analysis of the ONU in the Watchful Sleep Mode

With the above Markov chain model, we could obtain all of the steady-state probabilities under four different sleep period variation patterns. In this section, we define some indexes to evaluate the system performances via numerical analysis.

A. Energy-Saving Efficiency of the ONU

The power consumption of per unit scheduling cycle for an ONU when it stays in the states of *active held*, *active free*, *aware*, *listen*, and *sleep* were denoted as $W(\mathrm{Ah})$, $W(\mathrm{Af})$, $W(\mathrm{Aw})$, $W(L)$, and $W(S)$, respectively. We assumed that the power consumption of the Ah, Af, and Aw states follow the restriction of $W(\mathrm{Ah}) = W(\mathrm{Af}) = W(\mathrm{Aw}) = W(A)$ because these states are all FULL power states. Thus, the average energy consumption can be calculated as follows:

$$\begin{aligned}E[P_{sw}]=W(A)&\left[\sum_i P(\mathrm{Ah}_i)+P(\mathrm{Af})+\sum_{j=1}^{x}\mathrm{Aw}_j\right]\\&+W(L)\left[\sum_{k=1}^{n}\sum_{l=1}^{y}P(L_{k,l})\right]+W(S)\sum_{k=1}^{n}P(S_k).\end{aligned}\tag{15}$$

Hence, the energy-saving efficiency can be derived from the formula as follows:

$$\eta=\frac{W(A)-E[P_{sw}]}{W(A)}\times 100\%.\tag{16}$$

B. Average Data Packet Delay of the ONU

The average data packet delay (DPD) for those packets that arrive at the ONU, which stays in the Ah, Af, Aw, L, and S states, are, respectively, denoted as $E[d|\mathrm{Ah}]$, $E[d|\mathrm{Af}]$, $E[d|\mathrm{Aw}]$, $E[d|L]$, and $E[d|S]$. Thus, the total average DPD can be calculated as

$$\begin{aligned}E[D]=E[\delta|\mathrm{Ah}]&\cdot\sum_{i=1}^{\infty}P(\mathrm{Ah}_i)+E[\delta|\mathrm{Af}]\cdot P(\mathrm{Af})\\&+E[\delta|L]\cdot\sum_{k=1}^{n}\sum_{l=1}^{y}P(L_{k,l})+E[\delta|S]\cdot\sum_{k=1}^{n}P(\mathrm{S}_k)\\&+E[\delta|\mathrm{Aw}]\cdot\sum_{j=1}^{x}P(\mathrm{Aw}_j).\end{aligned}\tag{17}$$

Those packets that arrive when the ONU is sleeping have to wait until the ONU wakes up from the *sleep* state before being transmitted, because the duration time of the *sleep* state cannot be interrupted by the wakeup indication. In this case, the average state DPD consists of two parts: 1) the queuing time in the buffer; and 2) the "wait-to-wakeup" time when the ONU stays in the *sleep* state with z scheduling cycles, as expressed in Eq. (17). Here, the "wait-to-wakeup" time for a packet is $z/2$ cycles on average. For the other states except the *sleep* state, the average state DPD can just consider the queuing time because these states can be interrupted at any time by the wakeup indication, and the "wait-to-wakeup" time is almost zero. Therefore, in these cases, the state average DPD for a packet can be calculated in Eq. (18):

$$E[\delta|S] = z/2 \cdot T + \sum_{m=1}^{\infty} [p^a(i,z)(i-1) \cdot 1/\mu], \tag{18}$$

$$E[\delta|\mathrm{Ah}] = E[\delta|\mathrm{Af}] = E[\delta|\mathrm{Aw}] = E[\delta|L] = \sum_{m=1}^{\infty} [p^a(i,1)(i-1) \cdot 1/\mu]. \tag{19}$$

C. Normalized Cost Function of the ONU

In order to obtain the optimal trade-off between the above two conflicting performance indexes, we can derive the following normalized *cost* function with a similar concept as in [21] by using Eqs. (12) and (14). The normalized *cost* value means the expenditure, which should be paid to achieve the goal value P_g of the power consumption and the D_g of the data packet delay, which are the differences between the estimated values (i.e., $E[P_{sw}]$ and $E[D]$) and these goal values (i.e., P_g and D_g). The S_λ is the set of all the values of the arriving rate λ. By calculating the minimum *cost* for all cases in the set S_λ, the optimized key parameters can be obtained.

The weight coefficients of α and β take the value between 0 and 1 for two conflicting performance indexes (i.e., ESE and DPD), respectively, and follow the constraint of $\alpha + \beta = 1$. $(P_a - P_g)$ and $(D_{\max} - D_g)$ are used to normalize their estimated values, where $D_{\max}$ is the maximum DPD value for all cases of λ. In our simulations, the above two conflicting performance indexes are treated equally. Thus, the optimized parameters can be obtained by equating the two weights in Eq. (19), i.e., $\alpha = \beta = 0.5$:

$$\mathrm{Cost} = \left\{ \alpha \frac{\max[(E[P_{sw}] - P_g), 0]}{P_a - P_g} + \beta \frac{\max[(E[D] - D_g), 0]}{D_{\max} - D_g} \right\}. \tag{20}$$

IV. Performance Evaluation

This section evaluates the performances of the watchful sleep mode based on the Markov chain with the four different sleep period variation patterns. In the Markov chain model, we assumed the average service time of each packet is a scheduling cycle ($T_{\text{service}} = 2$ ms), which is $\mu = 1$. Thus, the value of the traffic arrival rate λ is equivalent to the value of the traffic load (i.e., λ/μ). The normalized power consumption of the ONU in the *active* (e.g., Ah, Af, Aw), *listen*, and *sleep* states are set to be 1, 0.4, 0.05, respectively [1].

In Subsection IV.A, we first give out the key performances (e.g., energy-saving efficiency and data packet delay) of the ONU under the different traffic arrival rate (λ) with the different values of the y and z parameters in Fig. 5. Then, comprehensive analyses are performed by considering various parameters (i.e., x, y, z, n) mentioned above in Section II, which are listed in Table I. In Subsection IV.B, we studied the impacts of the four different sleep period variation patterns on the normalized *cost* value of the ONU to obtain the optimal tradeoff between the above two conflicting performances. Note that all ONUs are assumed to have the same service level (e.g., in terms of traffic arrival rates) as in other research [17]. Therefore, the number of ONUs in a PON can be ignored in our simulations.

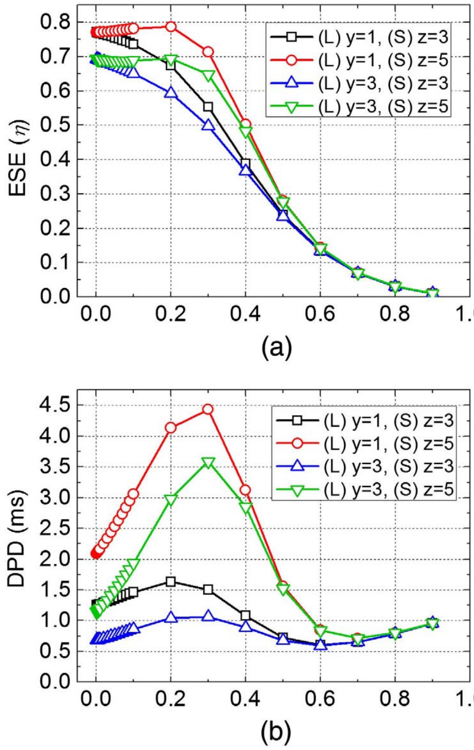

Fig. 5. Performance indexes of (a) ESE; (b) average DPD under the different traffic arrival rates (λ).

TABLE I
Notion of the System Parameters

Parameter	Notion
n	Number of (*sleep*, *listen*) state pairs
x	Number of substates of each *aware* state
y	Number of substates of each *listen* state
z	Number of scheduling cycles in each *sleep* state
λ	Traffic arrival rate

A. Performances Analysis Under the Constant Pattern

Figure 5 illustrates the performances of the ESE and the DPD under the different traffic rate λ, increasing from 0 to 1. The aware period (x) is fixed to be four scheduling cycles. In Fig. 5(a), when the traffic arrival rate λ is smaller, the higher ESE is obtained because the probabilities for both the *sleep* and *listen* states are larger. With increasing λ, the ESE decreases and converges to the same value. The smaller *listen* period and/or the larger *sleep* period tend to the higher ESE because the ONU has more of a chance to stay sleeping. In Fig. 5(b), the DPD presents a horizontal "S" curve, with λ increasing from 0 to 1. The state DPD is composed of the queuing delay and the "wait-to-wakeup" time, as shown in Eqs. (17) and (18). When λ is smaller, the DPD is dominated by the "wait-to-wakeup" time because the ONU is mostly in the *sleep* state. Hence, the DPD for $z = 5$ is much larger than that for $z = 3$. In the range of the light load, the queuing delay increases as λ increases, and consequently the DPD increases. When $\lambda = 0.3$ (for $z = 5$) and $\lambda = 0.2$ (for $z = 3$), the state DPD reaches the maximum. For the same z value, the larger y value means that the probability of entering the *sleep* state becomes smaller. As a result, the higher the y value is, the smaller the DPD is. When λ increases from 0.3 to 0.7 for $z = 5$ (from 0.2 to 0.6 for $z = 3$), the probability of the *sleep* state decreases, and the queuing delay turns out to be a dominant factor, which results in the DPD decreasing. When λ increases beyond a value (e.g., $\lambda = 0.7$), the DPD increases again, due to the growing importance of the queuing delay.

Figure 6 illustrates that the ESE and DPD would become larger, as the sleep period (z) increases, under the three cases of $\lambda = 0.2$, 0.5, and 0.8, with consideration of the different values of x and n, respectively. In the case of $\lambda = 0.2$, the probability of the *sleep* state is much larger, which not only results in higher ESE, but the effect of the value variations of x and n on the ESE and the state DPD can also be negligible (all curves are close together). When $\lambda = 0.5$ or 0.8, all the curves of the ESE and the state DPD are separated into four groups. The case is determined mainly by the x value and the effect of the n value would be minimal. The larger the x value is, the lower both the ESE and DPD are. This is due to the fact that the larger duration of the *aware* state would lead to higher power consumption. When $\lambda = 0.8$, the ESE and the DPD are below 0.05 and 3 ms. Figure 6 shows that the effect of the n value would be minimal. Therefore, the value of n is fixed to be 10 in the following simulations.

In Fig. 7, we study the impact of the different values of x, y, and z on the performances of the ESE and DPD, under $\lambda = 0.2$, 0.5, and 0.8, respectively. When λ is smaller (e.g., $\lambda = 0.2$), it is found that the effect of x on the ESE and DPD is almost negligible. But the performances of the ESE and DPD when $y = 1$ are slightly better than that in the case of $y = 3$. When $\lambda = 0.5$ or 0.8, the effect of y on the ESE and DPD is almost negligible, while the effect of x cannot be negligible.

B. Normalized Cost Comparison for Four Different Sleep Period Variation Patterns

In the above simulations, we just consider the constant pattern for sleep period variation; the z value will not change with the number n of state pairs. In this part, we will investigate the impact of the four different sleep period variation patterns on the normalized *cost* value, as defined in Eq. (19). We assume that the numbers of substates of *aware* (Aw) and *listen* (L) are 4 ($x = 4$) and 3

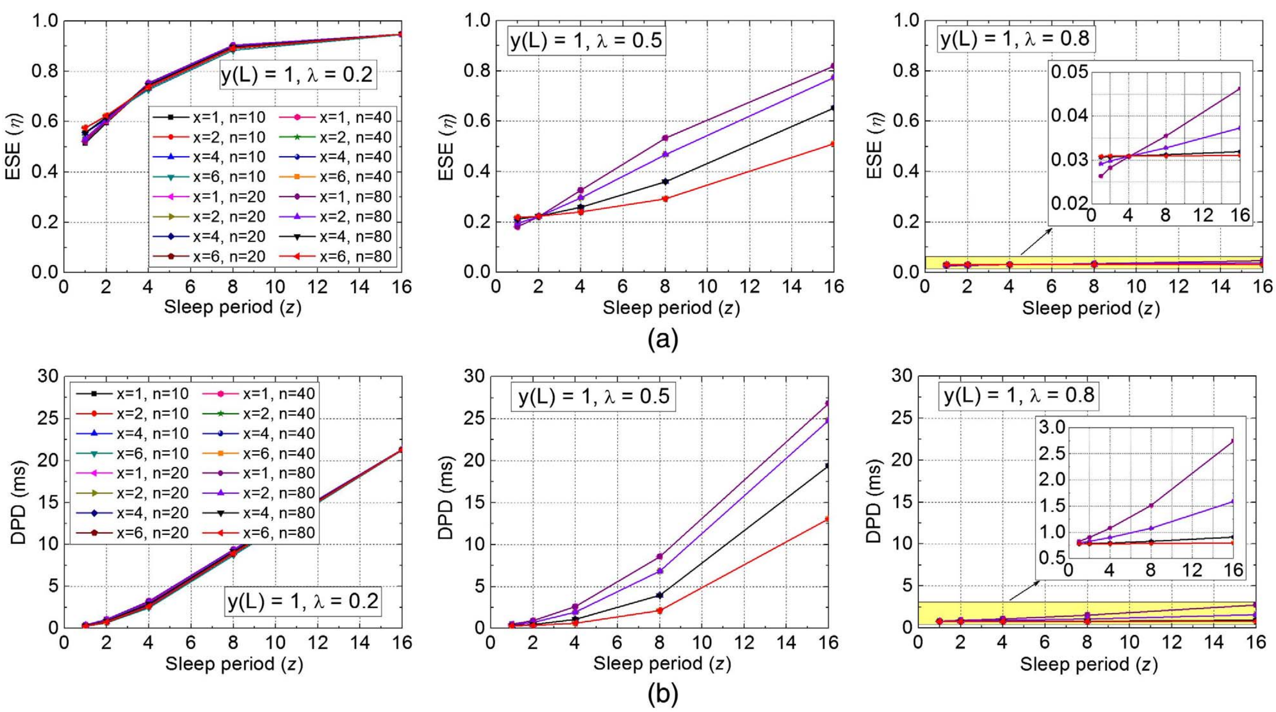

Fig. 6. Performances of (a) ESE and (b) average DPD versus the sleep period (z) with different x and n ($x = 1$, 2, 4, 6 and $n = 10$, 20, 40, 80).

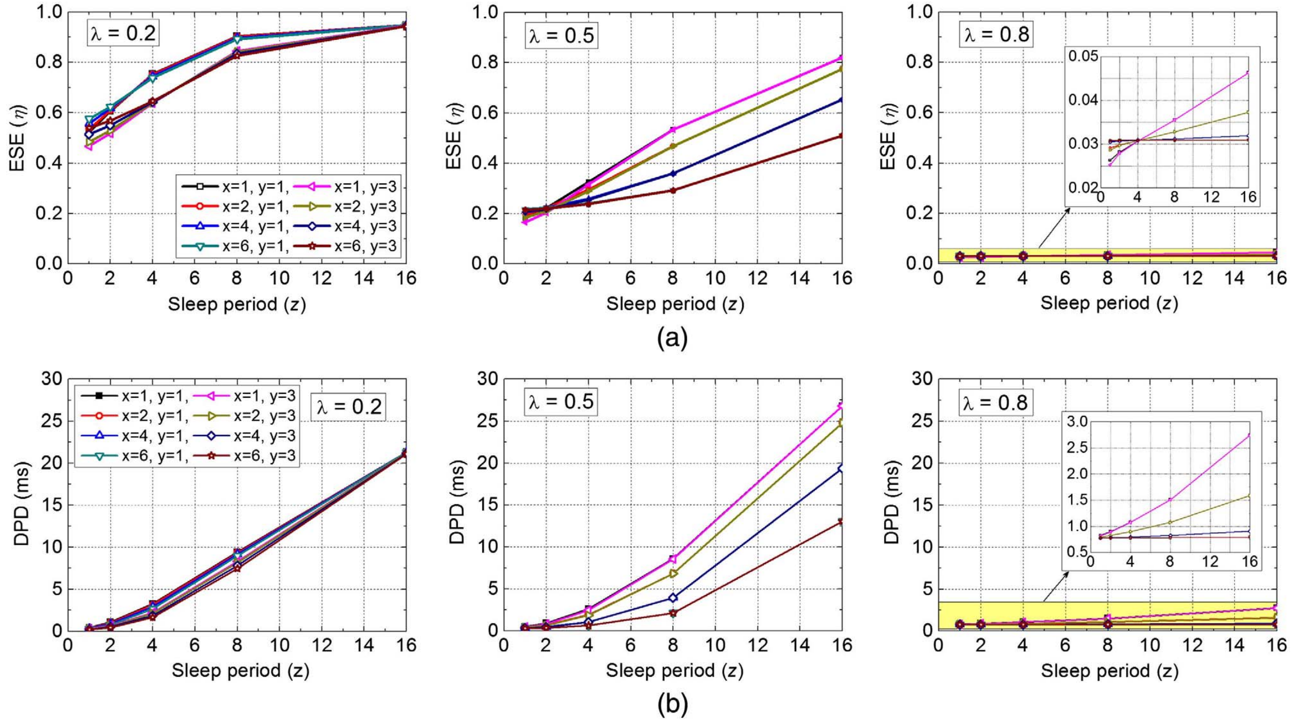

Fig. 7. Performances of (a) ESE and (b) average DPD versus the sleep period (z) with different values of x and y ($x = 1, 2, 4, 6$ and $y = 1, 3$).

($y = 3$), respectively. It is assumed that $\alpha = \beta = 0.5$ to equally treat the ESE and the state DPD. It is noted that the maximum value $D_{\max}$ of the DPD may differ depending on the number n of state pairs. The power goal value P_g is set to be $0.05P_a$. Under the *exponential* (i.e., $z = 2^{n-1}$), *linear*_2 (i.e., $z = 2n$), *linear*_1 (i.e., $z = n$), and constant (e.g., $z = 1$) patterns, we can find from Fig. 8 that the minimum *cost* values are 0.1505, 0.1499, 0.1457, and 0.1727, in the case of the $D_g = 2$ ms, respectively. When the $D_g = 4$ ms, the minimum *cost* values turn out to be 0.1505, 0.1268, 0.1406, and 0.1661, respectively. It is clear that the *linear*_2 pattern would be the best suitable one to achieve an optimal trade-off between the ESE and DPD performances.

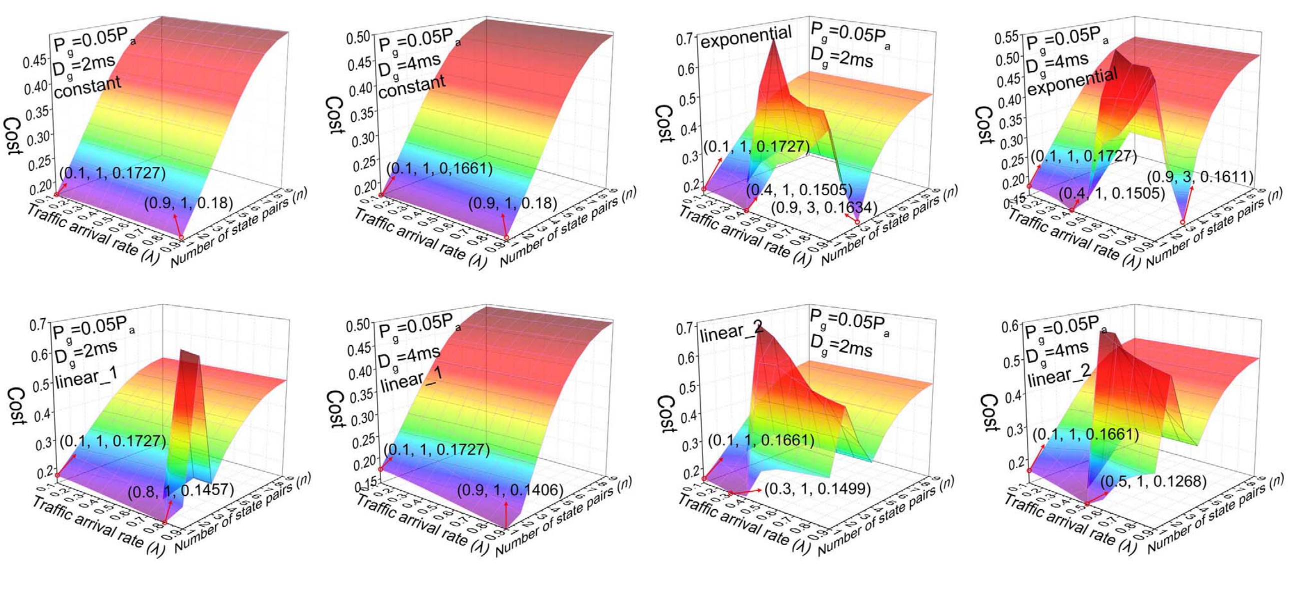

Fig. 8. Normalized *cost* value versus the traffic arrival rate (λ) and the number of state pairs (n) under four different sleep period variation patterns.

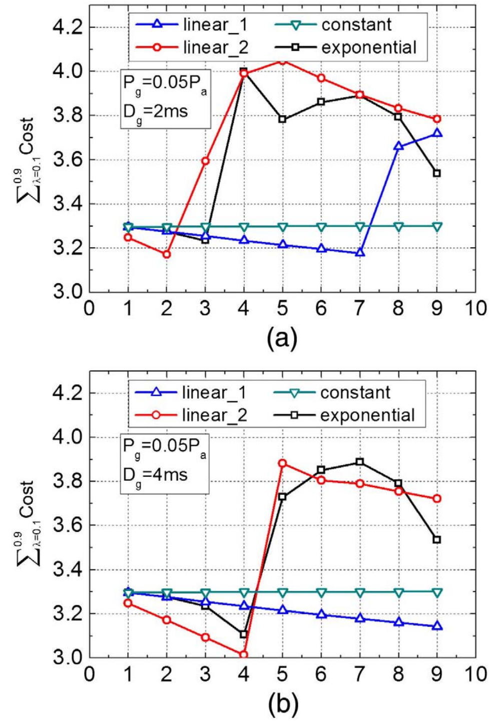

Fig. 9. Total sum of normalized *cost* values versus the number n of (*sleep*, *listen*) state pairs for the goals of $P_g = 0.05P_a$ and (a) $D_g = 2$ ms; (b) $D_g = 4$ ms.

To present the total sum of the normalized *cost* value for the different traffic arrival rates (λ), the simulation results are shown in Fig. 9. Under the constant pattern, the normalized *cost* value remains unchanged (e.g., *cost* value = 3) and is independent of the value n. In Fig. 8(a), the *linear*_1, *linear*_2, and *exponential* patterns can obtain optimal *cost* values of 3.2, 3.2, and 3.25, when the n is 7, 2, and 3, respectively, for the goals of $P_g = 0.05P_a$ and $D_g = 2$ ms. Figure 8(b) reveals that the *linear*_1, *linear*_2, and *exponential* patterns would reach optimal *cost* values of 3.2, 3.1, and 3.2, when n is 9, 4, and 4, respectively, for the goals of $P_g = 0.05P_a$ and $D_g = 4$ ms.

V. Conclusion

In this paper, we have modeled the watchful sleep mode based on the Markov chain model to analyze the performances in terms of the energy-saving efficiency and data packet delay. We also proposed four different sleep period variation patterns to reveal the contribution of the extended sleep periods to the performances in terms of the integrated performance in terms of the normalized *cost* value. Thereby, we can obtain an optimal trade-off between the above two conflicting performances in the different cases. From the extensive simulations, we found that (i) under the constant pattern, the effect of the number (n) of (*sleep*, *listen*) state pairs would be minimal for the watchful sleep mode. The number (z) of scheduling cycles in each *sleep* state is a critical parameter to the system performances. Under a light load, the effect of the parameter y is more important than that of the parameter x. Under the higher load, the effect of x cannot be negligible. (ii) By applying four different sleep period variation patterns, it is found that the *linear*_1, *linear*_2, and *exponential* patterns are better than the *constant* pattern. Those optimized *cost* values can be achieved when the parameter n is 7, 2, 3 and 9, 4, 4, and the delay goal value of D_g is 2 ms and 4 ms, respectively.

Acknowledgment

The work was jointly supported by the National Natural Science Foundation of China (NSFC) (61401087), the Natural Science Foundation of Jiangsu Province (BK20140642), the Fundamental Research Funds for the Central Universities of China (2242015R30011), and the Research Innovation for College Graduates of Jiangsu Province (SJLX16_0080).

References

[1] R. O. C. Hirafuji, K. B. da Cunha, D. R. Campelo, A. R. Dhaini, and D. A. Khotimsky, "The watchful sleep mode: A new standard for energy efficiency in the future access network," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 150–157, Aug. 2015.

[2] A. R. Dhaini, P. H. Ho, and G. Shen, "Toward green next-generation passive optical networks," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 94–101, Nov. 2011.

[3] B. Muzamil, M. Sohail, J. K. Kasi, and A. K. Kasi, "Performance analysis of passive optical networks with energy saving through the integrated sleep mode," *Opt. Switching Netw.*, vol. 21, pp. 16–30, 2016.

[4] G. J. Nasr, "Climate change and environment in the Arab world," Lecture delivered at AUB, Issam Fares Institute for Public Policy and International Affairs, 2009.

[5] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: Overview and reduction strategies," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 62–69, June 2011.

[6] K. Hinton, J. Baliga, M. Feng, R. Ayre, and R. S. Tucker, "Power consumption and energy efficiency in the Internet," *IEEE Netw.*, vol. 25, no. 2, pp. 6–12, Mar./Apr. 2011.

[7] L. Nuaymi, *WiMAX: Technology for Broadband Wireless Access*. West Sussex, England: Wiley, 2008.

[8] C. Lange, M. Braune, and N. Gieschen, "On the energy consumption of FTTB and FTTH access networks," in *Optical Fiber Communication Conf. and the Nat. Fiber Optic Engineers Conf.*, San Diego, 2008, paper JWA105.

[9] J. Zhang, M. T. Hosseinabadi, and N. Ansari, "Standards-compliant EPON sleep control for energy efficiency: Design and analysis," *J. Opt. Commun. Netw.*, vol. 5, no. 7, pp. 677–685, 2013.

[10] P. Vetter, D. Suvakovic, H. Chow, P. Anthapadmanabhan, K. Kanonakis, K.-L. Lee, F. Saliou, X. Yin, and B. Lannoo, "Energy-efficiency improvements for optical access," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 136–144, 2014.

[11] B. Muzamil and P. Saengudomlert, "Integrated sleep mode for improving energy efficiency of NG-PONs," in *11th Int.*

Conf. on Optical Communications and Networks (ICOCN), 2012, pp. 1–4.

[12] H. S. Abbas and M. A. Gregory, "The next generation of passive optical networks: A review," *J. Netw. Comput. Appl.*, vol. 67, pp. 53–74, 2016.

[13] L. Valcarenghi, D. P. Van, P. G. Raponi, P. Castoldi, D. R. Campelo, S.-W. Wong, S.-H. Yen, L. G. Kazovsky, and S. Yamashita, "Energy efficiency in passive optical networks: Where, when, and how?" *IEEE Netw.*, vol. 26, no. 6, pp. 61–68, 2012.

[14] "Gigabit-capable passive optical network (XG-PON): Transmission convergence (TC) specification," Release 1, ITU-T G.987.3, 2010.

[15] M. Zhu, X. Zeng, Y. Lin, and X. Sun, "Modeling and analysis for watchful sleep mode in PON power management," in *15th Int. Conf. on Optical Communications and Networks (ICOCN)*, 2016.

[16] D. A. Khotimsky, D. Zhang, L. Yuan, R. O. C. Hirafuji, and D. R. Campelo, "Unifying sleep and doze modes for energy-efficient PON systems," *IEEE Commun. Lett.*, vol. 18, no. 4, pp. 688–691, Apr. 2014.

[17] H. Bang, J. Kim, S. Lee, and C. Park, "Determination of sleep period for cyclic sleep mode in XG-PON power management," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 98–100, Jan. 2012.

[18] M. Zhu, X. Zeng, Y. Lin, and X. Sun, "Impact of sleep period variation pattern on energy-delay performances of watchful sleep mode in PON," in *Asia Communications and Photonics Conf.*, 2016.

[19] M. Frydenberg, "The chain graph Markov property," *Scand. J. Stat.*, vol. 17, pp. 333–353, 1990.

[20] K. Lange, "Markov chain Monte Carlo methods," in *Mathematical and Statistical Methods for Genetic Analysis*. New York: Springer, 1997, pp. 142–163.

[21] W. Shen and Q.-A. Zeng, "Cost-function-based network selection strategy in integrated wireless and mobile network," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3778–3788, Nov. 2008.