数理解析研究所講究録
1132巻 2000年 125-133

Computational Aspect of the Power–Effective Auto Sleep Scheduling for a Buffer System

岡村寛之 †, 土肥正 †, 尾崎俊治 †
HIROYUKI OKAMURA†, TADASHI DOHI† and SHUNJI OSAKI†
†Department of Industrial and Systems Engineering, Hiroshima University
4–1 Kagamiyama 1 Chome, Higashi-Hiroshima 739–8527, Japan.

Abstract

This paper addresses a problem of how to determine the optimal auto sleep time when the computer user should turn the hard disk or the display to a sleep mode in order to save the electrical power after the computer has not been accessed. We propose a stochastic model to obtain the optimal sleep timing strategy which maximizes the power effectiveness, where access requirements arrive at the system according to a renewal process and are processed by a general service time. Then the phase-type approximations are proposed to represent the power effectiveness. We investigate the approximation performance for the proposed methods through a simulation study.

1 Introduction

Recently, the automatic sleep function of the hard disk or the display in a computer system is rapidly recognized to be important in terms of power management. In fact, the auto sleep function is equipped in almost computer systems as a standard function. Then, the optimal design for the auto sleep function is the most important problem, in particular, for notebook computers with limited capacity of battery. For example, on the hard disk of a computer, the electrical power consumed to warm up from sleep mode is larger than that consumed in the normal operation. Thus, it is not always effective to design the system such that moves its state to the sleep mode whenever there is no access requirement.

First, the optimal design problem for the auto sleep function was considered by Sandoh, Hirakoshi and Kawai [1]. Dohi, Kaio and Osaki [2] proposed a statistical nonparametric method to estimate the optimal sleep timing for the same problem. However, it is noted that the seminal works above simplified the underlying problem extremely and was incomplete for representation of stochastic behavior of the auto sleep system. More valid formulations were made by Okamura, Dohi and Osaki [3, 4]. They considered two kinds of models (Type I model and Type II model) with and without cancellation of access requirements arrived at the system, respectively. More specifically, Type I model with cancellation assumes that other access requirements arrived at the system when one job has been processed are canceled, and focuses on the multi-use circumstance for a desktop computer unit. On the other hand, Type II model corresponds to a buffer system in which the other access requirements are accumulated while one job has been processed, and deals with the multi-job system such as network printers. They proved that the optimal sleep timing strategies for both models are *the switching strategies*, *i.e.*, turn always the system to a sleep mode after the process for a job is completed, or not do at all, if the access requirements arrive according to the homogeneous Poisson process.

However, if the access requirements arrive following more general stochastic processes such as the renewal process, it is difficult to obtain the power effectiveness explicitly.

Okamura, Dohi and Osaki [3, 4] applied the simple parametric approximation methods by Miyazawa [5] and the usual diffusion approximation to represent the power effectiveness, but could not obtain the satisfactory approximation performance. The main reason for this problem is that the arrival process may belong to a more wide class of stochastic processes. In this paper, we apply the phase-type approximations to represent the power effectiveness and derive the approximated optimal auto sleep time for Type II model. Altiok [6] and Heijden [7] showed that the phase-type approximations are useful to represent the general probability distributions. Asmussen and Koole [11] have also proved that the phase-type renewal process is weakly dense in the class of stationary simple point processes.

The paper is planed as follows. Section 2 describes the auto sleep model under consideration and gives an implicit form of the power effectiveness under the assumption that access requirements arrive at the system following the renewal process. Section 3 concerns the approximation problem for the power effectiveness. Then the phase-type approximation is introduced to represent the access requirements process. Furthermore, two estimation methods for the phase-type approximation are proposed. Section 4 is devoted to investigate the approximation performance for the proposed methods through a simulation study. Finally, the paper is concluded with some remarks.

2 Model Description

2.1 Notation and assumptions

Suppose that the access requirements arrive at the system according to an ordinary renewal process $\{N(t);\ t > 0\}$. Denote a sequence of inter-arrival times between $(k-1)$-th and k-th arrivals by $\{X_k;\ k = 1, 2, \cdots\}$. Then, X_k are the non-negative i.i.d. random variables, having the probability distribution $F(t)$ with mean $1/\lambda$ (> 0) and variance σ_a (> 0). The tasks required by the k-th access are processed with the times S_k, which are the non-negative i.i.d. random variables having the probability distribution $H(t)$ with finite mean $1/\mu$ (> 0) and variance σ_s (> 0). It is assumed that the system under consideration can take the following states;

Busy: The system processes some tasks required by accesses, where the set-up time τ is needed before processing each task. After the present task is completed, the state of system moves to the idle state. During the busy state, the electrical power consumed per unit time is P_1 (> 0).

Idle: No access requirement occurs, after one task is completed. If a new access requirement occurs until the total spent time in the idle period becomes t_0, the system begins to process it after elapsing τ time units. Otherwise, the state of system moves to the sleep state at the moment when the total spent time in the idle period becomes t_0. Throughout this paper, we call t_0 *the auto sleep time*. The electrical power consumed per unit time during the idle period is also P_1 (> 0).

Sleep: The sleep state is the lower-power state, so that the electrical power consumed per unit time is less than that in the other states. To simplify the discussion, we assume that the electrical power consumed per unit time in the sleep state is zero. When an access requirement occurs, the sleep mode terminates immediately and the state of system moves to the warm-up state.

Warm-up: In order to begin processing a task from the sleep mode, s (> 0) time units are needed for warming-up. Hence, after $s + \tau$ time units are elapsed, the process

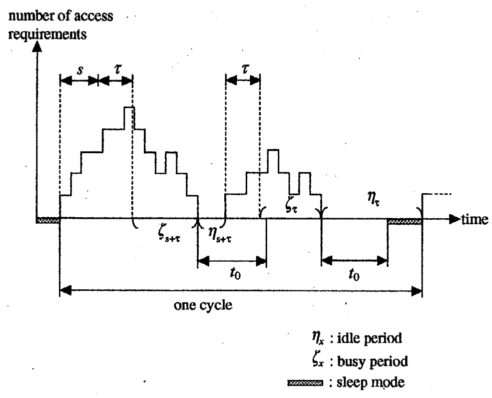

Figure 1: Possible realization of the stochastic system.

for the task is started. In the warm-up state, the electrical power P_2 is consumed per unit time, where $P_2 > P_1$.

In this paper, it is assumed that the other access requirements arrived while the system is in the busy are accumulated in an infinite buffer. Hence, the state of system moves to an idle after all tasks accumulated in the buffer are completed. Figure 1 is depicted a possible realization of the system. Since access requirements are processed according to FIFO (first-in-first-out) discipline, the number of access requirements in the buffer forms a $GI/G/1/FIFO$ queueing process.

2.2 Formulation of the power effectiveness

Let us formulate the power effectiveness criterion. The power effectiveness means the time measure when the system is operative by unit electrical power, so that the mean operative time per unit electrical power. In this model, the system starts processing after a delayed time period $s+\tau$ or τ. Such a delayed time period is called the *vacation*, which pays a significant role in the application of queueing system. In the queueing system with vacation, the periods when the buffer is empty, when the server processes tasks and when the system is in the vacation are called idle, busy and dormant periods, respectively. It is well known that the following relation holds among the idle, busy and dormant periods;

$$\text{E[busy period]} = \rho\Big\{\text{E[dormant period]} + \text{E[busy period]} + \text{E[idle period]}\Big\}, \quad (1)$$

where ρ is the traffic intensity and is defined by $\rho = \lambda/\mu$.

Now define the following random variables;

η_t: time length when the system is in the idle state, having the distribution function $I(x|t) = \Pr\{\eta_t \le x\}$ and the survivor function $\overline{I}(x|t) = 1 - I(x|t)$,

ζ_t: time length when the system is in the busy state,

where both subscripts t of the random variables above indicate the delayed time periods, so that $t = s+\tau$ or $t = \tau$. Then we have, from Eq. (1),

$$\frac{\mathrm{E}[\zeta_t]}{t + \mathrm{E}[\zeta_t] + \mathrm{E}[\eta_t]} = \rho, \quad (t = \tau, s+\tau). \quad (2)$$

We define the time period from the beginning of warm-up state to the next beginning of that as one cycle. The mean operating time during one cycle is

$$A(t_0) = \frac{1}{1-\rho}\Big\{\tau + \mathrm{E}[\eta_{s+\tau}] + \mathrm{E}[N](\tau + \mathrm{E}[\eta_\tau])\Big\}, \tag{3}$$

where

$$\mathrm{E}[N] = \frac{\Pr\{\eta_{s+\tau} \le t_0\}}{\Pr\{\eta_\tau > t_0\}} = \frac{I(t_0 \mid s+\tau)}{\overline{I}(t_0 \mid \tau)}. \tag{4}$$

Also, the expected power consumed for one cycle is

$$\begin{aligned} C(t_0) &= \Big\{\frac{\rho}{1-\rho}P_1 + P_2\Big\}s + \frac{P_1\tau}{1-\rho} + P_1\Big\{\frac{\rho}{1-\rho}\mathrm{E}[\eta_{s+\tau}] + \mathrm{E}[\eta_{s+\tau} \wedge t_0]\Big\} \\ &\quad + \mathrm{E}[N]\Big\{\frac{P_1\tau}{1-\rho} + P_1\Big(\frac{\rho}{1-\rho}\mathrm{E}[\eta_\tau] + \mathrm{E}[\eta_\tau \wedge t_0]\Big)\Big\}, \end{aligned} \tag{5}$$

where $\mathrm{E}[\eta_t \wedge t_0] = \mathrm{E}[\min(\eta_t, t_0)] = \int_0^{t_0} u dI(u \mid t) + t_0\overline{I}(t_0 \mid t)$. From the renewal reward theorem, the power effectiveness is given by

$$\begin{aligned} W(t_0) &= \lim_{t\to\infty} \frac{\mathrm{E}[\text{the operative time length in } (0,t]]}{\mathrm{E}[\text{the total power consumed in } (0,t]]} \\ &= A(t_0)/C(t_0). \end{aligned} \tag{6}$$

The problem is to find the optimal auto sleep time t_0^* which maximizes the power effectiveness $W(t_0)$, that is, $\max_{0\le t_0<\infty} W(t_0)$.

3 The phase-type approximation

In general arrival cases, it is difficult to obtain an explicit forms of $I(t|x)$ and the power effectiveness $W(t_0)$. Hence any approximation method has to be developed to generate the auto sleep time in the computer operation phase. From these motivations, we introduce the phase-type approximation method.

3.1 Formulation of the power effectiveness based on the phase-type approximation

Consider a Markov process on the state space $\{1, 2, \cdots, m+1\}$, where $\{1, 2, \cdots, m\}$ denote the transient states and $\{m+1\}$ means the absorbing one. The initial probability vector for the Markov process is given by $(\boldsymbol{\alpha}, 0)$. Until the absorption in the state $m+1$, the process behaves similar to the Markov process with an infinitesimal generator T, where T is a matrix with components λ_{ij} (> 0), $1 \le i, j \le m$, $j \ne i$ and $-\lambda_{ii}$ (< 0). In our model, the absorption implies the occurrence of events, such as the arrival of access requirements, *etc.* After the absorption, the process is restarted with initial state. Then, the inter-arrival time distribution is assumed to obey the following phase-type distribution with parameter $(\boldsymbol{\alpha}, T)$;

$$F_{PH}(t) = 1 - \boldsymbol{\alpha}\exp(Tt)\boldsymbol{e}, \tag{7}$$

where $\boldsymbol{e}$ is a column vector of 1s.

Denote N_t and J_t be the number of arrivals in $(0, t]$ and the internal state of arrival at time t, respectively, where the internal states can be interpreted as the states of various factors which cause the arrival of access requirements. We define the transition probability

$$P_{ij}(n, t) = \Pr\{N_t = n, J_t = j \mid N_0 = 0, J_0 = i\} \tag{8}$$

and the matrix $P(n,t)$ with components $P_{ij}(n,t)$. The matrix generating function $P^*(z,t)$ is obtained as follows.

$$P^*(z,t) = \sum_{n=0}^{\infty} P(n,t)z^n = \exp\{(T + z\boldsymbol{T^0}\boldsymbol{\alpha})t\}, \tag{9}$$

where $\boldsymbol{T^0} = -T\boldsymbol{e}$ is the column vector. We can find that both the number of access requirements N_t and the internal state process J_t construct an embedded Markov chain at the points of the completion of a process. Let A_n and B_n be the $m \times m$ matrices with components $[A_n]_{ij}$ and $[B_n]_{ij}$, respectively, for $n \geq 0$. The component $[A_n]_{ij}$ is the probability that the internal state moves from i to j and that n access requirements during a process do not occur. Thus, it is easy to obtain

$$A_n = \int_0^{\infty} P(n,t)dH(t). \tag{10}$$

The component $[B_n]_{ij}$ is also the probability of a transition from the internal state i to j and that n access requirements in the buffer are remained. Hence, the matrix B_n is

$$B_n = \int_0^{\infty} \boldsymbol{e}\boldsymbol{\alpha} \sum_{k=0}^{n} P(k,x)P(n-k,t)dH(t), \tag{11}$$

where x is the server vacation period.

Denote $\boldsymbol{g}$ be the probability vector with components g_i, $1 \leq i \leq m$, where g_i is the probability that the internal state is in i at the beginning of the idle period. Then, the probability distribution of the idle period can be reduced to the phase-type distribution with parameter $(\boldsymbol{g}, T)$. Therefore, we concentrate our attention to find the probability vector $\boldsymbol{g}$. Consider the first-passage time when the number of access requirements independent of the internal state becomes i from $i+1$, where such a time is called *the fundamental period.* Define the $m \times m$ matrix G as the transition probability matrix with components $[G]_{ij}$, which is the probability that the internal state moves from i to j during the fundamental period. We also define the $m \times m$ matrix K as the transition probability matrix for the first-passage time when the number of access requirements becomes 0 from 0 again. Note that the matrix K is constructed with the probability that the internal state moves from i to j during the fundamental period. By the above definitions, we can see that the following equations hold (see Lucantoni, Meier-Hellstern and Neuts [8]).

$$G = \sum_{n=0}^{\infty} A_n G^n, \quad K = \sum_{n=0}^{\infty} B_n G^n \quad \text{and} \quad \boldsymbol{g}K = \boldsymbol{g}. \tag{12}$$

The algorithm for computation of the matrix G was proposed by Lucantoni and Ramaswami [10]. Equations (9), (11) and (12) yield

$$\boldsymbol{g} = \frac{\boldsymbol{\alpha}\exp\{(T + \boldsymbol{T^0}\boldsymbol{\alpha}G)x\}\, G}{\boldsymbol{\alpha}\exp\{(T + \boldsymbol{T^0}\boldsymbol{\alpha}G)x\}\, G\boldsymbol{e}}. \tag{13}$$

In the sequel, the probability distribution of the idle period can be approximated by

$$I(t|x) \approx 1 - \boldsymbol{g}\exp(Tt)\boldsymbol{e}, \tag{14}$$

which results an approximation form of the power effectiveness.

3.2 Statistical estimation procedure

Since the phase-type renewal process is composed of two stochastic processes which are observable and unobservable, usual statistical estimation methods, such as the method of maximum likelihood, cannot be used for model parameters. Thus, we introduce the following two estimation methods for the arrival process.

(1) The moment matching

Heijden [7] proposed the following moment matching conditions. If there are n unknown-parameters, they are determined by fitting the first n moments to the sample moments estimated from real data. If the inter-arrival time distribution of the phase-type renewal process obeys the following Coxian-2 distribution;

$$T = \begin{bmatrix} -\lambda_1 & 0 \\ \lambda_2 & -\lambda_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\alpha} = (1-a, a), \tag{15}$$

then the estimators for the parameters are given by

$$a = \frac{\lambda_2}{\lambda_1}(m_1\lambda - 1), \tag{16}$$

$$\lambda_1 = \frac{3m_1m_2 - m_3 - \sqrt{m_3^2 + 18m_2^3 + 24m_1^3m_3 - 9m_1m_2(3m_1m_2 + 2m_3)}}{3m_2^2 - 2m_1m_3} \tag{17}$$

and

$$\lambda_2 = \frac{2(m_1\lambda_1 - 1)}{m_2\lambda_1 - 2m_1}, \tag{18}$$

where m_1, m_2 and m_3 are the first three moments of inter-arrival time.

(2) The EM-algorithm for phase-type distribution

The EM (expectation-maximization) algorithm is an iterative method for maximum likelihood estimation [12, 13]. It is a useful methods to parameterize statistical models including the incomplete data. Suppose that $Y = u(X)$ is observed and that X is unobserved, where Y and X have the probability density functions g_γ and f_γ, respectively. Then $(n+1)$-th step in the EM algorithm is to find the value γ_{n+1} which maximizes

$$\gamma \to \mathrm{E}[\log f_\gamma(X) \mid u(X) = y;\, \gamma_n], \tag{19}$$

where y is the observed data and γ_n is the current estimate after n steps of the algorithm (see *e.g.* [14] for detail). In particular, when the inter-arrival time distribution has the phase-type distribution, the EM-algorithm is given by as follows:

Let $(y_1, y_2, \ldots, y_n)$ be the observed sample data. Then $(k+1)$-th iteration of the algorithm becomes

E-Step: Calculate

$$\pi_i^{(k+1)} = \sum_{l=1}^{n} \mathrm{E}[\pi_i^{(k)} | y_l;\, \hat{\boldsymbol{\alpha}}^{(k)}, \hat{T}^{(k)}] \quad \text{for } i = 1, \ldots, m, \tag{20}$$

$$\xi_i^{(k+1)} = \sum_{l=1}^{n} \mathrm{E}[\xi_i^{(k)} | y_l;\, \hat{\boldsymbol{\alpha}}^{(k)}, \hat{T}^{(k)}] \quad \text{for } i = 1, \ldots, m, \tag{21}$$

$$\Lambda_{ij}^{(k+1)} = \sum_{l=1}^{n} \mathrm{E}[\Lambda_{ij}^{(k)} | y_l;\, \hat{\boldsymbol{\alpha}}^{(k)}, \hat{T}^{(k)}] \quad \text{for } i \neq j, i = 1, \ldots, m \text{ and } j = 1, \ldots, m \tag{22}$$

M-Step: The new estimators are given by

$$\hat{\alpha}_i^{(k+1)} = \frac{\pi_i^{(k+1)}}{n}, \quad \hat{t}_{ij}^{(k+1)} = \frac{\Lambda_{ij}^{(k+1)}}{\xi_i^{(k+1)}}, \quad \hat{t}_{ii}^{(k+1)} = -\left(\frac{\Lambda_{i0}^{(k+1)}}{\xi_i^{(k+1)}} + \sum_{j=1,j\neq 1}^{m} \hat{t}_{ij}^{(k+1)}\right), \tag{23}$$

where $\hat{\alpha}_i$ and $\hat{t}_{ij}$ are the elements of $\hat{\boldsymbol{\alpha}}$ and $\hat{T}$, respectively. In the above expressions, π_i is the number of Markov processes starting in state i, ξ_i is the total time spent in state i and Λ_{ij} is the total number of jumps from state i to j.

4 Numerical Examples

In this section, we investigate the approximation performance of the phase-type methods proposed in Section 3. Suppose that the arrival of access requirements follows the renewal process with the Weibull inter-arrival time distribution;

$$F(t) = 1 - \exp\{-(t/\beta_a)^{m_a}\}, \quad (m_a = 0.5,\ \beta_a = \rho/\Gamma(1 + 1/m_a)) \tag{24}$$

where $m_a\,(> 0)$ and $\beta_a\,(> 0)$ denote the shape and scale parameters of the Weibull distribution, respectively, and where $\Gamma(\cdot)$ is the standard gamma function. We also suppose that the processing time distribution is the exponential distribution; $H(t) = 1 - \exp(-t)$. The other model parameters are fixed as $P_1 = 1.0$, $P_2 = 3.0$, $\tau = 0.1$ and $s = 5.0$. In our approximation scheme, the inter-arrival time distribution of the phase-type renewal process corresponds to the Coxian-2 distribution. In addition to the phase-type approximation, we calculate the optimal auto sleep time based on the equilibrium approximation [4] and compare their precision, where the equilibrium approximation is to approximate the idle period distribution with the equilibrium distribution of inter-arrival time, that is

$$I(t|x) \approx F_e(t) = \lambda \int_0^t \overline{F}(x)dx. \tag{25}$$

Tables 1 and 2 present the optimal auto sleep times and their associated maximum power effectiveness based on the equilibrium approximation and the phase-type approximations. In the phase-type approximations, we use the moment matching and the EM-algorithm to estimate the model parameters. In addition, we estimate numerically the power effectiveness by the Monte Carlo simulation, provided that the auto sleep time is given. On each table, the values in brackets indicate the lower and upper bounds on the confidence interval with significant level 95% and are calculated by the simulation. From Table 2, it is observed that the maximum power effectiveness by the moment matching is not belonging to the corresponding confidence intervals. This result shows that the phase-type approximation with the moment matching may not function well to approximate the power effectiveness. Also, the power effectiveness estimated by the simulation are smaller than those based on the other approximations. On the other hand, the simulated power effectiveness based on the equilibrium approximation and the phase-type approximation with the EM-algorithm tend to belong to the confidence interval in the heavy traffic circumstance. Comparing Table 1 with Table 2, the simulated power effectiveness when the optimal auto sleep time is calculated by the EM-algorithm is larger than the simulated one by the equilibrium approximation. Hence we can conclude that the actual optimal auto sleep time is close to that estimated by the EM-algorithm, and that it outperforms the equilibrium approximation, when the system has a heavy traffic intensity. However, in the case of the light traffic circumstance, it is seen that the optimal auto sleep time based on

Table 1: The optimal auto sleep time based on the equilibrium approximation.

ρ	$\hat{t}_0^*$	$\hat{W}(t_0^*)$	$W(\hat{t}_0^*)$
0.1	0.00	0.266	0.162 (0.154, 0.171)
0.2	0.58	0.326	0.235 (0.223, 0.247)
0.3	1.69	0.393	0.286 (0.268, 0.304)
0.4	2.74	0.467	0.410 (0.380, 0.438)
0.5	3.74	0.548	0.468 (0.441, 0.495)
0.6	4.75	0.633	0.623 (0.572, 0.675)
0.7	5.83	0.722	0.715 (0.664, 0.766)
0.8	7.01	0.814	0.802 (0.759, 0.845)
0.9	8.43	0.908	0.850 (0.798, 0.901)

Table 2: The optimal auto sleep time based on the phase-type approximation.

	moment matching			EM-algorithm		
ρ	$\hat{t}_0^*$	$\hat{W}(t_0^*)$	$W(\hat{t}_0^*)$	$\hat{t}_0^*$	$\hat{W}(t_0^*)$	$W(\hat{t}_0^*)$
0.1	0.00	0.184	0.161 (0.153, 0.170)	0.00	0.159	0.161 (0.153, 0.170)
0.2	1.84	0.277	0.229 (0.214, 0.243)	0.00	0.292	0.226 (0.214, 0.238)
0.3	1.38	0.371	0.291 (0.275, 0.307)	0.00	0.268	0.283 (0.265, 0.301)
0.4	0.83	0.464	0.357 (0.335, 0.380)	0.28	0.417	0.352 (0.332, 0.373)
0.5	0.46	0.556	0.483 (0.450, 0.516)	∞	0.519	0.567 (0.529, 0.606)
0.6	0.24	0.644	0.556 (0.505, 0.607)	∞	0.612	0.638 (0.585, 0.691)
0.7	0.13	0.730	0.628 (0.584, 0.672)	∞	0.710	0.697 (0.652, 0.743)
0.8	0.10	0.815	0.740 (0.680, 0.801)	∞	0.805	0.829 (0.774, 0.884)
0.9	0.12	0.905	0.822 (0.759, 0.886)	∞	0.907	0.879 (0.834, 0.925)

the equilibrium approximation tends to give the lager power effectiveness. It follows from these results that the phase-type approximation with the EM-algorithm is more efficient in the cases of the heavy traffic intensity.

5 Concluding Remarks

In this paper, we have considered the stochastic auto sleep model under the renewal arrival process, and have proposed phase-type approximation methods to represent the power effectiveness. Based on these approximations, we have calculated the optimal auto sleep schedule which maximizes the power effectiveness. In numerical examples, we have investigated the approximation performance for the proposed methods. As a result, we have shown that the phase-type approximation could be useful for finding the optimal auto sleep time approximately in the heavy traffic circumstance for the other approximation method.

References

[1] Sandoh, H., Hirakoshi, H. and Kawai, H. (1996), An optimal time to sleep for an auto-sleep system, *Computers & Operations Research*, **23**, 221–227.

[2] Dohi, T., Kaio, N. and Osaki, S. (1997), Nonparametric approach to power-saving strategies for a portable personal computer, *Electronics and Communications in Japan, Part 3*, **80**, 80–90.

[3] Okamura, H., Dohi, T. and Osaki, S. (1999), The Auto-sleep scheduling based on power effectiveness for a computer system (in Japanese), *Transactions on IEICE*, **J82-A**, 1808–1818.

[4] Okamura, H., Dohi, T. and Osaki, S. (2000), A power-effective design of auto-sleep mode for computer hard disks (in Japanese), to appear in *Transactions on SICE.*

[5] Miyazawa, M. (1989), A generalized Pollaczek-Khintchine formula for the $GI/GI/1/K$ queue and its application to approximation, *Stochastic Models*, **3**, 53–65.

[6] Altiok, T. (1985), On the phase-type approximations of general distributions , *IIE Transactions*, **17**, 110–116.

[7] Heijden, M. C. (1988), On the three-moment approximation of a general distribution by a Coxian distribution, *Probability in the Engineering and Informational Sciences*, **2**, 257–261.

[8] Lucantoni, D. M., Meier-Hellstern, K. S. and Neuts, M. F. (1990), A single-server queue with server vacations and a class of non-renewal arrival processes, *Advances in Applied Probability*, **22**, 676–705.

[9] Takagi, H. (1991), *Queueing Analysis, Vol. 1*, Elsevier Science Publisher, Netherlands.

[10] Lucantoni, D. M. and Ramaswami, V. (1985), Efficient algorithms for solving the non-linear matrix equations arising in phase type queues, *Stochastic Models*, **1**, 29–51.

[11] Asmussen, S. and Koole, G. (1993), Marked point processes as limits of Markovian arrival streams, *Journal of Applied Probability*, **30**, 356–372.

[12] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser. B*, **39**, 1–38.

[13] Wu, C. F. J. (1983), On the convergence properties of the EM algorithm, *Ann. Statist.*, **11**, 95–103.

[14] Asmussen, S., Nerman, O. and Olsson, M. (1996), Fitting phase-type distributions via the EM algorithm, *Scand. J. Statist.*, **23**, 419–441.