

コンピュータシステムの自動スリープ機能における省電力効果

岡村寛之, 土肥正, 尾崎俊治
広島大学工学部第二類 (電気系)
東広島市鏡山 1-4-1
tel: 0824-24-7697
fax: 0824-22-7025
e-mail: dohi@gal.sys.hiroshima-u.ac.jp

あらまし 近年, コンピュータシステムにおいて, ハードディスクやディスプレイの自動スリープ機能が省電力の観点から重要視されるようになってきた. 本稿では, コンピュータへのアクセス要求がない瞬間において, 省電力の観点からユーザが如何にハードディスクをスリープ状態に切り替えるかについての問題を取り上げる. ここでは特に, 2種類の確率モデルを提案する. 最初のモデルにおいて, コンピュータへのアクセス要求が到着した後, その処理期間中では他のアクセス要求はキャンセルされる. 一方, 2番目のモデルでは, 他のアクセス要求がバッファ内で待ち行列を形成することを許容している. アクセス要求がマルコフ的に到着する場合, 定常状態における電力有効性を最大にする最適自動スリープ戦略が導出される. より一般的な到着パターンに対しては, いくつかの近似解が提供される.

On the Effects of Power Saving in Auto-Sleep Functions on A Computer System

H. Okamura, T. Dohi and S. Osaki
Hiroshima University, Dept. of Industrial and Systems Engineering,
4-1 Kagamiyama 1 Chome, Higashi-Hiroshima 739, Japan.
tel: +81-824-24-7697
fax: +81-824-22-7025
e-mail: dohi@gal.sys.hiroshima-u.ac.jp

Abstract Recently, the automatic sleep function of the hard disk or the display in a computer system is rapidly recognized to be important in order to save the electric power. This paper addresses a problem of how to determine the optimal timing when the user should turn the hard disk off in terms of power saving if the computer has not been accessed. We propose two stochastic models to analyse the optimal timing strategy. In the first model, the access requirement arrives at the computer, but the other requirements are cancelled during the processing time interval. On the other hand, the second model permits the possibility that the other requirements form a queue in the buffer. We derive the optimal auto-sleep strategy which maximizes the so-called power effectiveness criterion for the case of the Markovian arrival, and provide some approximated solutions for more general cases.

1. はじめに

通常、コンピュータの省電力化対策として、CPU、液晶画面のバックライト、ハードディスク装置など、電力消費の比較的大きい部分に対して自動スリープ機能が搭載されている。コンピュータの自動スリープ機能とは、ある一定時間（停止時間）以上コンピュータに対するアクセス要求がない場合、自動的にコンピュータを一時停止（スリープ）させる機能であり、停止時間は使用者が任意に設定することが可能である。しかしながら、その決定は使用者の勘と経験によって決定されているのが現状であり、省電力という観点からその妥当性は明らかにされていない。

そこで文献 [1, 2] では、コンピュータシステムの自動スリープ機能を (i) コンピュータへのアクセス要求は確率的かつ断続的に発生する、(ii) コンピュータに一定期間 t_0 以上アクセス要求がない場合、コンピュータは自動的にスリープモードに入るという仮定の下で、ハードディスクの省電力方策の観点から最適な停止時間を導出している。しかしながら、文献 [1, 2] において設定された、アクセス中に他のアクセス要求は生起しないという仮定は、より一般的な自動スリープ機能を考える場合必ずしも妥当であるとは限らず、結果をそのまま適用することができない。

そこで本稿では、アクセスの処理中に他のアクセス要求が発生することを許容し、そのアクセス要求が (i) キャンセルされる場合、(ii) 待ち行列を形成する場合の 2 種類の自動スリープ機能に対して考察を行う。特にアクセスの発生がポアソン過程に従う場合には、電力有効性を最大にするような最適停止時間が解析的に導出される。

2. コンピュータシステムの自動スリープ機能 (モデル 1)

文献 [1, 2] で定義されたコンピュータシステムの自動スリープ機能を考える。ある一定時間以上コンピュータへのアクセス要求がなければ、その時点からコンピュータを一時停止（スリープ）するようなスリープ機能を考える。アクセス要求はある確率過程に従って生起するものとし、アクセス要求が発生してから実際にアクセスに対する処理を開始するまでの準備時間を τ (定数) とする。単一のアクセスの処理に要する時間は確率分布 $G(t)$ に従う独立で同

一な確率変数 S であり、その平均と分散を、それぞれ $E[S] = 1/\mu$, $\text{Var}[S] = \sigma_s^2$ で表す。また、アクセスの処理中に発生した他のアクセス要求はキャンセルされるものとし、アクセスの処理が完了してからコンピュータが待機状態である時間が t_0 に達する場合、コンピュータはその時点からスリープモードに入る。システムがスリープ状態の時にアクセス要求が発生すると、コンピュータの起動時間 s (定数) を経過してシステムは稼働状態になる。このような時間 t_0 を停止時間と呼ぶ (図 1 参照)。システムが

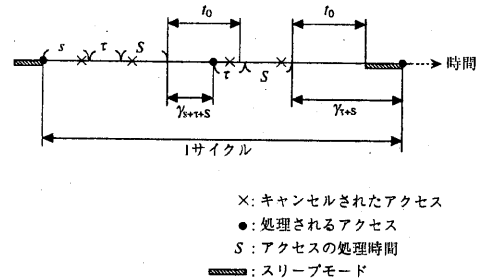


図 1: 自動スリープ機能の概念図 (モデル 1)。

稼働状態または待機状態であるとき、単位時間当たりに消費する通常電力を P_1 、システムを起動する際に単位時間当たりに消費する電力を P_2 ($P_2 > P_1$)、システムがスリープモードにあるとき電力の消費はないものとする。

(2.1) 再生過程

アクセス要求の発生時間間隔 X_k ($k = 1, 2, \dots$) は互いに独立で同一な確率分布関数 $F(t)$ に従い、その平均を $E[X_k] = 1/\lambda$ 、分散を $\text{Var}[X_k] = \sigma_a^2$ とする。特に、システムを起動してから、最初のスリープモードが終了するまでの期間を 1 サイクルとする。アクセスの処理が完了してから時間 t_0 中にアクセス要求が発生するという事象が、1 サイクル中に起こる回数を N とする。また、アクセス要求が時刻 0 で発生したとき、時刻 t より後に初めてアクセス要求が発生するまでの時間を表す確率変数 (残存寿命) を γ_t とすると、 N は 1 回目の成功する確率が $\Pr\{\gamma_{\tau+s} \leq t_0\}$ 、2 回目以降で成功する確率が $\Pr\{\gamma_{\tau+s} \leq t_0\}$ の幾何分布に従うため、 N の期待値は $E[N] = \Pr\{\gamma_{\tau+s} \leq t_0\} / \Pr\{\gamma_{\tau+s} > t_0\}$

となる。よって1サイクルの稼働時間は

$$\begin{aligned} A^1(t_0) &= \tau + 1/\mu + E[N](\tau + 1/\mu) \\ &= \tau + 1/\mu + [\tau + 1/\mu] \\ &\quad \times \frac{\int_0^\infty H(t_0|s + \tau + x)dG(x)}{\int_0^\infty \bar{H}(t_0|\tau + x)dG(x)} \quad (1) \end{aligned}$$

となる。ここで $H(x|t)$ は γ_t の確率分布関数で、 $\bar{H}(\cdot|t) = 1 - H(\cdot|t)$ である。これは、 X_k に関する再生関数 $M(t) = \sum_{n=1}^\infty F^{(n)}(t)$ ($F^{(n)}(t)$ は $F(t)$ の n 重畳み込み) を用いて

$$H(x|t) = F(t+x) - \int_0^t \bar{F}(t+x-y)dM(y), \quad (2)$$

$$(\bar{F}(\cdot) = 1 - F(\cdot))$$

$$E[\gamma_t] = \frac{1}{\lambda}[1 + M(t)] - t \quad (3)$$

のように表される。また、1サイクルにおける総期待電力は

$$\begin{aligned} C^1(t_0) &= P_2 s + P_1(\tau + 1/\mu) \\ &\quad + P_1 \int_0^\infty E[\gamma_{s+\tau+x} \wedge t_0]dG(x) \\ &\quad + [P_1(\tau + 1/\mu) \\ &\quad + P_1 \int_0^\infty E[\gamma_{\tau+x} \wedge t_0]dG(x)] \\ &\quad \times \frac{\int_0^\infty H(t_0|s + \tau + x)dG(x)}{\int_0^\infty \bar{H}(t_0|\tau + x)dG(x)}. \quad (4) \end{aligned}$$

ここで、

$$\begin{aligned} E[\gamma_t \wedge t_0] &= E[\min(\gamma_t, t_0)] \\ &= \int_0^{t_0} u dH(u|t) + t_0 \bar{H}(t_0|t) \quad (5) \end{aligned}$$

である。これより、単位電力で、何時間システムを稼働させられるかを示す電力有効性は

$$W^1(t_0) = \frac{A^1(t_0)}{C^1(t_0)} \quad (6)$$

となる。

(2.2) ポアソン過程

ここでは、再生過程の特別な場合として、アクセス要求の発生がパラメータ $\lambda (> 0)$ のポアソン過程に従う場合を考える。このとき再生関数は $M(t) = \lambda t$ となるので、以下に示す式が成立する。

$$H(x|t) = 1 - e^{-\lambda x}, \quad (7)$$

$$E[\gamma_t] = 1/\lambda, \quad (8)$$

$$E[\gamma_t \wedge t_0] = (1/\lambda)(1 - e^{-\lambda t_0}). \quad (9)$$

よって、1サイクルの稼働時間は

$$A_{\text{poi}}^1(t_0) = \frac{1}{e^{-\lambda t_0}}(\tau + 1/\mu), \quad (10)$$

1サイクルにおける総期待電力は

$$C_{\text{poi}}^1(t_0) = \frac{P_1}{e^{-\lambda t_0}} \left[\tau + 1/\mu + \frac{1 - e^{-\lambda t_0}}{\lambda} \right] + P_2 s, \quad (11)$$

電力有効性は

$$W_{\text{poi}}^1(t_0) = A_{\text{poi}}^1(t_0)/C_{\text{poi}}^1(t_0) \quad (12)$$

となる。このとき、電力有効性 $W_{\text{poi}}(t_0)$ を最大にする最適停止時間 t_0^* に関して以下の定理を得る。

定理 1: $P_2/P_1 < 1/[\lambda s]$ ならば、電力有効性 $W_{\text{poi}}(t_0)$ を最大にする最適停止時間は $t_0^* = 0$ 、すなわちアクセスの処理が完了したと同時にスリープモードに入ることが最適となり、そのときの最大電力有効性は $W_{\text{poi}}(0)$ となる。逆に $P_2/P_1 \geq 1/[\lambda s]$ ならば $t_0^* \rightarrow \infty$ 、すなわち全くスリープモードに入らないことが最適となり、そのときの最大電力有効性は $W_{\text{poi}}(\infty)$ となる。□

(2.3) 近似モデル

一般の再生過程に対して残存寿命分布 $H(x|t)$ を求めるためには、再生関数 $M(t)$ を特定する必要がある。しかしながら、再生関数を求めることは一般的に困難であることに注意しなければならない。そこで、残存寿命分布 $H(x|t)$ の近似表現を用いることにより、(2.1) で導出された電力有効性 $W^1(t_0)$ を近似する。

ここでは、Miyazawa [3] によって提案された以下の2つの近似を考える。

$$\begin{aligned} \text{近似 1: } H(x|t) &\approx H_{\text{APP1}}(x) \\ &= F_e(x), \quad (13) \end{aligned}$$

$$\begin{aligned} \text{近似 2: } H(x|t) &\approx H_{\text{APP2}}(x) \\ &= 1 - \frac{\mu \int_0^\infty e^{-\mu u} \bar{F}(x+u) du}{1 - \bar{F}(\mu)}. \quad (14) \end{aligned}$$

ここで

$$F_e(x) = \lambda \int_0^x \bar{F}(u) du, \quad (15)$$

$$\bar{F}(x) = \int_0^\infty e^{-x u} dF(u). \quad (16)$$

近似1は式(2)が定常状態において式(15)の均衡分布に一致するという性質から得られたものであり、近似2は処理時間分布に指数分布を仮定し、 $X_k > S$ という条件の下で $X_k - S$ を残存寿命として近似したものである。これより、以下のような期待残存寿命に関する諸量を表現することが出来る。

$$\begin{aligned} \text{近似1: } E[\gamma_t] &\approx E[\gamma_{\text{APP1}}] \\ &= \frac{\lambda^2 \sigma_a^2 + 1}{2\lambda}, \end{aligned} \quad (17)$$

$$\begin{aligned} E[\gamma_t \wedge t_0] &\approx E[\gamma_{\text{APP1}} \wedge t_0] \\ &= \int_0^{t_0} \int_x^\infty \bar{F}(u) du dx, \end{aligned} \quad (18)$$

$$\begin{aligned} \text{近似2: } E[\gamma_t] &\approx E[\gamma_{\text{APP2}}] \\ &= \frac{1}{\lambda[1 - \bar{F}(\mu)]} - \frac{1}{\mu}, \end{aligned} \quad (19)$$

$$\begin{aligned} E[\gamma_t \wedge t_0] &\approx E[\gamma_{\text{APP2}} \wedge t_0] \\ &= \mu \int_0^{t_0} \int_0^\infty \frac{\bar{F}(x+u)}{1 - \bar{F}(\mu)} du dx. \end{aligned} \quad (20)$$

$H_{\text{APP1}}(x)$ と $H_{\text{APP2}}(x)$ が t に依存しないことに注意して、式(1)と(4)に式(17)-(20)を代入すると、1サイクルの稼働時間は

$$A_{\text{APP1}}^1(t_0) = \frac{1}{\bar{H}_{\text{APP1}}(t_0)} (\tau + 1/\mu), \quad (21)$$

$$A_{\text{APP2}}^1(t_0) = \frac{1}{\bar{H}_{\text{APP2}}(t_0)} (\tau + 1/\mu) \quad (22)$$

となり、1サイクルにおける総期待電力は

$$C_{\text{APP1}}^1(t_0) = \frac{P_1}{\bar{H}_{\text{APP1}}(t_0)} (\tau + 1/\mu + E[\gamma_{\text{APP1}} \wedge t_0]) + P_2 s, \quad (23)$$

$$C_{\text{APP2}}^1(t_0) = \frac{P_1}{\bar{H}_{\text{APP2}}(t_0)} (\tau + 1/\mu + E[\gamma_{\text{APP2}} \wedge t_0]) + P_2 s \quad (24)$$

となる。これらを用いて、電力有効性の近似式

$$W_{\text{APP1}}^1(t_0) = A_{\text{APP1}}^1(t_0) / C_{\text{APP1}}^1(t_0), \quad (25)$$

$$W_{\text{APP2}}^1(t_0) = A_{\text{APP2}}^1(t_0) / C_{\text{APP2}}^1(t_0) \quad (26)$$

が得られる。さらに、 $H_{\text{APP1}}(t)$ と $H_{\text{APP2}}(t)$ のハザード率、 $\tau_{\text{APP1}}(t)$ および $\tau_{\text{APP2}}(t)$ は次のようになる。

$$\tau_{\text{APP1}}(t) = \frac{\bar{F}(t)}{\int_t^\infty \bar{F}(u) du}, \quad (27)$$

$$\tau_{\text{APP2}}(t) = \frac{e^{-\mu t} \bar{F}(t)}{\int_t^\infty e^{-\mu u} \bar{F}(u) du} - \mu. \quad (28)$$

以上より、 $W_{\text{APP1}}^1(t_0)$ 、ならびに $W_{\text{APP2}}^1(t_0)$ を最大にする最適停止時間 t_0^* に関する定理が得られる。

定理2 (近似1):

1) アクセス時間間隔の確率分布関数 $F(t)$ が狭義DHR (Decreasing Hazard Rate) とする。

i) $1/[\lambda s] < P_2/P_1 < 1/[\tau_{\text{APP1}}(\infty)s]$ ならば、電力有効性を最大にする有限で唯一の最適停止時間 $t_0^* \in (0, \infty)$ が存在し、

$$P_1 = P_2 \tau_{\text{APP1}}(t_0^*) s \quad (29)$$

を満足する。

ii) $P_2/P_1 \leq 1/[\lambda s]$ ならば最適停止時間は $t_0^* = 0$ となる。

iii) $P_2/P_1 \geq 1/[\tau_{\text{APP1}}(\infty)s]$ ならば最適停止時間は $t_0^* \rightarrow \infty$ となる。

2) アクセス時間間隔の確率分布関数 $F(t)$ がIHR (Increasing Hazard Rate) とする。

i) $P_2/P_1 < E[\gamma_{\text{APP1}}]/s$ ならば $t_0^* = 0$,

ii) $P_2/P_1 \geq E[\gamma_{\text{APP1}}]/s$ ならば $t_0^* \rightarrow \infty$. \square

定理3 (近似2):

1) アクセス時間間隔の確率分布関数 $F(t)$ が狭義DHR とする。

i) $1/[\tau_{\text{APP2}}(0)s] < P_2/P_1 < 1/[\tau_{\text{APP2}}(\infty)s]$ ならば、電力有効性を最大にする有限で唯一の最適停止時間 $t_0^* \in (0, \infty)$ が存在し、

$$P_1 = P_2 \tau_{\text{APP2}}(t_0^*) s \quad (30)$$

を満足する。

ii) $P_2/P_1 \leq 1/[\tau_{\text{APP2}}(0)s]$ ならば最適停止時間は $t_0^* = 0$ となる。

iii) $P_2/P_1 \geq 1/[\tau_{\text{APP2}}(\infty)s]$ ならば最適停止時間は $t_0^* \rightarrow \infty$ となる。

2) アクセス時間間隔の確率分布関数 $F(t)$ がIHR (Increasing Hazard Rate) とする。

- i) $P_2/P_1 < E[\gamma_{APP2}]/s$ ならば $t_0^* = 0$,
 ii) $P_2/P_1 \geq E[\gamma_{APP2}]/s$ ならば $t_0^* \rightarrow \infty$. \square

3. コンピュータシステムの自動スリープ機能 (モデル 2)

2. ではアクセスの処理中に発生する他のアクセス要求はキャンセルされると仮定したが, ここではアクセスの処理中に発生した他のアクセス要求が待ち行列を形成するモデルについて考える.

(3.1) GI/G/1 過程

アクセスの発生時間間隔 $X_k (k = 1, 2, \dots)$ が互いに独立で同一な確率分布関数 $F(t)$ ($E[X_k] = 1/\lambda$, $\text{Var}[X_k] = \sigma_a^2$) に従い, 2. と同様にアクセスの処理に要する時間が平均 $1/\mu (> 0)$ の確率変数 S である場合を考える. このとき, アクセスの発生は再生過程となるため, システム内の挙動は $GI/G/1$ 待ち行列を形成する (図 2 参照). 一般性を失うこ

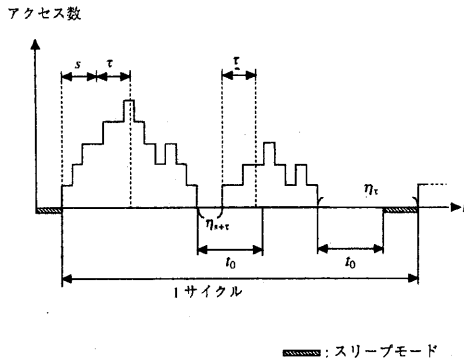


図 2: 自動スリープ機能の概念図 (モデル 2).

となく, $GI/G/1$ 待ち行列のトラフィック強度に関して $\rho = \lambda/\mu < 1$ を仮定する. モデル 1 と同様に, システムを起動してから最初のスリープモードが終了するまでの期間を 1 サイクルとすると, 1 サイクルの稼動時間は

$$A^2(t_0) = \frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} (s + E[\eta_{s+\tau}]) + \left(\frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\eta_\tau] \right) \frac{I(t_0|s+\tau)}{I(t_0|\tau)} \quad (31)$$

ここで, η_t は $GI/G/1$ 待ち行列の余暇期間が t 時間である時のサーバの空き期間を表す確率変数であ

り, $I(x|t)$ はその確率分布関数である. また, 1 サイクルにおける総期待電力は

$$C^2(t_0) = \left(\frac{\rho}{1-\rho} P_1 + P_2 \right) s + \frac{P_1 \tau}{1-\rho} + P_1 \left(\frac{\rho}{1-\rho} E[\eta_{s+\tau}] + E[\eta_{s+\tau} \wedge t_0] \right) + \left[\frac{P_1 \tau}{1-\rho} + P_1 \left(\frac{\rho}{1-\rho} E[\eta_\tau] + E[\eta_\tau \wedge t_0] \right) \right] \frac{I(t_0|s+\tau)}{I(t_0|\tau)} \quad (32)$$

となる. これより, 電力有効性は

$$W^2(t_0) = \frac{A^2(t_0)}{C^2(t_0)} \quad (33)$$

となる.

(3.2) M/G/1 過程

ここでは, (3.1) の特別な場合としてアクセス要求の発生がパラメータ $\lambda (> 0)$ のポアソン過程 $\{N(t), t \geq 0\}$ に従う場合を考える. このとき, システム内のアクセス要求は $M/G/1$ 待ち行列を形成し, 以下に示す式が成立する.

$$I(x|t) = 1 - e^{-\lambda x}, \quad (34)$$

$$E[\eta_t] = 1/\lambda, \quad (35)$$

$$E[\eta_t \wedge t_0] = (1/\lambda)(1 - e^{-\lambda t_0}). \quad (36)$$

よって 1 サイクルの稼動時間は

$$A_{M/G/1}^2(t_0) = \frac{1}{e^{-\lambda t_0}} \left(\frac{\tau}{1-\rho} + \frac{\rho}{\lambda(1-\rho)} \right) + \frac{\rho}{1-\rho} s. \quad (37)$$

1 サイクルにおける総期待電力は

$$C_{M/G/1}^2(t_0) = \frac{P_1}{(1-\rho)e^{-\lambda t_0}} \times \left[\tau + \frac{1}{\lambda} \{1 - (1-\rho)e^{-\lambda t_0}\} \right] + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s \quad (38)$$

となる. これより, 電力有効性は

$$W_{M/G/1}^2(t_0) = A_{M/G/1}^2(t_0)/C_{M/G/1}^2(t_0) \quad (39)$$

となる. このとき, 電力有効性 $W_{M/G/1}^2(t_0)$ を最大にする最適停止時間 t_0^* に関して, 条件 $\rho < 1$ の下で次の結果が得られる.

定理 4: 条件 $\rho < 1$ を仮定する. このとき,

$$(i) \frac{P_2}{P_1} < \frac{1}{\lambda s} + \frac{\rho}{\rho + \lambda \tau} \Rightarrow t_0^* = 0,$$

$$(ii) \frac{P_2}{P_1} \geq \frac{1}{\lambda s} + \frac{\rho}{\rho + \lambda \tau} \Rightarrow t_0^* \rightarrow \infty. \quad \square$$

(3.3) Miyazawa [3] の方法による近似モデル

GI/G/1 待ち行列の空き期間 η_t の確率分布関数 $I(x|t)$ を求めることは非常に困難である。そこで、 $I(x|t)$ を近似することにより、(3.1) で導出された電力有効性 $W^2(t_0)$ を近似する。

(2.3) と同様に 2 つの近似式を適用する。

$$\begin{aligned} \text{近似 1: } I(x|t) &\approx I_{APP1}(x) \\ &= F_e(x), \end{aligned} \quad (40)$$

$$\begin{aligned} \text{近似 2: } I(x|t) &\approx I_{APP2}(x) \\ &= 1 - \frac{\mu \int_0^\infty e^{-\mu u} \bar{F}(x+u) du}{1 - \bar{F}(\mu)}. \end{aligned} \quad (41)$$

この近似表現を用いると、1 サイクルの稼働時間は

$$\begin{aligned} A_{APP1}^2(t_0) &= \frac{1}{\bar{I}_{APP1}(t_0)} \left(\frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\eta_{APP1}] \right) \\ &\quad + \frac{\rho}{1-\rho} s, \end{aligned} \quad (42)$$

$$\begin{aligned} A_{APP2}^2(t_0) &= \frac{1}{\bar{I}_{APP2}(t_0)} \left(\frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\eta_{APP2}] \right) \\ &\quad + \frac{\rho}{1-\rho} s \end{aligned} \quad (43)$$

となり、1 サイクルにおける総期待電力は

$$\begin{aligned} C_{APP1}^2(t_0) &= \frac{P_1}{\bar{I}_{APP1}(t_0)} \left(\frac{\tau}{1-\rho} \right. \\ &\quad \left. + \frac{\rho}{1-\rho} E[\eta_{APP1}] + E[\eta_{APP1} \wedge t_0] \right) \\ &\quad + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s, \end{aligned} \quad (44)$$

$$\begin{aligned} C_{APP2}^2(t_0) &= \frac{P_1}{\bar{I}_{APP2}(t_0)} \left(\frac{\tau}{1-\rho} \right. \\ &\quad \left. + \frac{\rho}{1-\rho} E[\eta_{APP2}] + E[\eta_{APP2} \wedge t_0] \right) \\ &\quad + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s \end{aligned} \quad (45)$$

となる。これらを用いて、電力有効性の近似式

$$W_{APP1}^2(t_0) = A_{APP1}^2(t_0)/C_{APP1}^2(t_0), \quad (46)$$

$$W_{APP2}^2(t_0) = A_{APP2}^2(t_0)/C_{APP2}^2(t_0) \quad (47)$$

が得られる。さらに、スイッチング関数

$$q_{APP1}^2(t_0) = [\tau_{APP1}(t_0)s\{\rho P_1 + (1-\rho)P_2\}$$

$$\begin{aligned} &- (1-\rho)P_1] \bar{I}_{APP1}(t_0) A_{APP1}^2(t_0) \\ &- \rho \tau_{APP1}(t_0) s \bar{I}_{APP1}(t_0) C_{APP1}^2(t_0), \end{aligned} \quad (48)$$

$$\begin{aligned} q_{APP2}^2(t_0) &= [\tau_{APP2}(t_0)s\{\rho P_1 + (1-\rho)P_2\} \\ &- (1-\rho)P_1] \bar{I}_{APP2}(t_0) A_{APP2}^2(t_0) \\ &- \rho \tau_{APP2}(t_0) s \bar{I}_{APP2}(t_0) C_{APP2}^2(t_0) \end{aligned} \quad (49)$$

を定義すると、 $W_{APP1}^2(t_0)$ ならびに $W_{APP2}^2(t_0)$ を最大にする最適停止時間 t_0^* に関する定理が得られる。

定理 5 (近似 1): 仮定 $\rho/[\rho P_1 + (1-\rho)P_2] < W_{APP1}^2(t_0)$ の下で、

1) アクセス時間間隔の確率分布関数 $F(t)$ が狭義 DHR とする。

i) $q_{APP1}^2(0) > 0$ かつ $q_{APP1}^2(\infty) < 0$ ならば、定常状態における単位時間当たりの期待電力を最小にする有限で唯一の最適停止時間 $t_0^* \in (0, \infty)$ が存在し、 $q_{APP1}^2(t_0^*) = 0$ を満足する。そのときの最大電力有効性は、

$$\begin{aligned} W_{APP1}^2(t_0^*) &= \rho \tau_{APP1}(t_0^*) s / \left[\rho P_1 \right. \\ &\quad \left. + (1-\rho)P_2 \right] \tau_{APP1}(t_0^*) s \\ &\quad - (1-\rho)P_1 \end{aligned} \quad (50)$$

となる。

ii) $q_{APP1}^2(0) \leq 0$ ならば最適停止時間は $t_0^* = 0$ となり、最大電力有効性は $W_{APP1}^2(0)$ となる。

iii) $q_{APP1}^2(\infty) \geq 0$ ならば最適停止時間は $t_0^* \rightarrow \infty$ となり、最大電力有効性は $W_{APP1}^2(\infty)$ となる。

2) アクセス時間間隔の確率分布関数 $F(t)$ が IHR とする。このとき

$$\frac{P_2}{P_1} < \frac{E[\eta_{APP1}]}{s} + \frac{\rho E[\eta_{APP1}]}{\tau + \rho E[\eta_{APP1}]} \quad (51)$$

ならば $t_0^* = 0$ となり、そうでなければ $t_0^* \rightarrow \infty$ となる。 \square

定理 6 (近似 2): 仮定 $\rho/[\rho P_1 + (1-\rho)P_2] < W_{APP2}^2(t_0)$ の下で、

1) アクセス時間間隔の確率分布関数 $F(t)$ が狭義 DHR とする。

i) $q_{APP2}^2(0) > 0$ かつ $q_{APP2}^2(\infty) < 0$ ならば、定常状態における単位時間当たりの期待電力を最小にする有限で唯一の最適停止時間 $t_0^* \in (0, \infty)$ が存在し、 $q_{APP2}^2(t_0^*) = 0$ を満足する。そのときの最大電力有効性は、

$$W_{APP2}^2(t_0^*) = \rho r_{APP2}(t_0^*)s / \left[\rho P_1 + (1 - \rho)P_2 \right] r_{APP2}(t_0^*)s - (1 - \rho)P_1 \quad (52)$$

となる。

ii) $q_{APP2}^2(0) \leq 0$ ならば最適停止時間は $t_0^* = 0$ となり、最大電力有効性は $W_{APP2}^2(0)$ となる。

iii) $q_{APP2}^2(\infty) \geq 0$ ならば最適停止時間は $t_0^* \rightarrow \infty$ となり、最大電力有効性は $W_{APP2}^2(\infty)$ となる。

2) アクセス時間間隔の確率分布関数 $F(t)$ が IHR とする。このとき

$$\frac{P_2}{P_1} < \frac{E[\eta_{APP2}]}{s} + \frac{\rho E[\eta_{APP2}]}{\tau + \rho E[\eta_{APP2}]} \quad (53)$$

ならば $t_0^* = 0$ となり、そうでなければ $t_0^* \rightarrow \infty$ となる。□

(3.4) 拡散近似モデル

ここでは、拡散近似と呼ばれる近似方法を用いて電力有効性 $W^2(t_0)$ を近似し、さらに近似的な最適解の導出を試みる。

アクセス要求の発生数の累積過程と $GI/G/1$ 過程を、それぞれ拡散過程 $\{W_i(t); t \geq 0\}$ ($i = 1, 2$) で近似する。拡散過程 $W_i(\cdot)$ ($i = 1, 2$) に対して無限小平平均 a_i と拡散パラメータ $b_i (> 0)$ を以下のように定義する。

$$a_i = \lim_{\Delta t \rightarrow 0} \frac{E[W_i(t + \Delta t) - W_i(t) | W_i(t)]}{\Delta t},$$

$$b_i = \lim_{\Delta t \rightarrow 0} \frac{\text{Var}[W_i(t + \Delta t) - W_i(t) | W_i(t)]}{\Delta t} \quad (i = 1, 2) \quad (54)$$

文献 [4, 5] の結果を直接用いることによって、

$$a_1 = \lambda (> 0), \quad b_1 = \lambda^3 \sigma_a^2,$$

$$a_2 = \lambda - \mu (< 0), \quad b_2 = \lambda^3 \sigma_a^2 + \mu^3 \sigma_s^2$$

を得る。また、拡散過程 $W_1(t)$ が初めて 1 に到達する時間を

$$\theta = \inf\{t \geq 0; W_1(t) \geq 1 | W_1(0) = 0\} \quad (55)$$

のように定義する。さらに、空き期間の確率分布関数 $I(x|t)$ を θ の確率分布関数で近似する。すなわち、

$$I(x|t) \approx I_{\text{dif}}(x) = \Pr\{\theta \leq x\} = \Phi\left[\frac{a_1 x - 1}{\sqrt{b_1 x}}\right] + e^{\frac{2a_1}{b_1}} \Phi\left[\frac{-a_1 x - 1}{\sqrt{b_1 x}}\right] \quad (56)$$

のような逆ガウス分布を考える。ここで、 $\Phi[\cdot]$ は標準正規分布関数である。また η_t に関連した諸量の期待値は

$$E[\eta_t] \approx E[\theta] = 1/a_1, \quad (57)$$

$$E[\eta_t \wedge t_0] \approx E[\theta \wedge t_0] = \frac{1}{a_1} \Phi\left[\frac{-1 + a_1 t_0}{\sqrt{b_1 t_0}}\right] + t_0 \Phi\left[\frac{1 + a_1 t_0}{\sqrt{b_1 t_0}}\right] - \left(\frac{1}{a_1} + t_0\right) e^{2a_1/b_1} \Phi\left[\frac{-1 - a_1 t_0}{\sqrt{b_1 t_0}}\right] \quad (58)$$

となる。これらを式 (31)-(33) に代入することによって、1 サイクルの稼働時間の近似式 $A_{\text{dif}}^2(t_0)$ 、1 サイクルにおける総期待電力の近似式 $C_{\text{dif}}^2(t_0)$ 、および電力有効性の近似式 $W_{\text{dif}}^2(t_0)$ が得られる。ここで、 $W_{\text{dif}}^2(t_0)$ を最大にする最適停止時間 t_0^* を得るために、以下に示すスイッチング関数 $q_{\text{dif}}^2(t_0)$ を定義する。

$$q_{\text{dif}}^2(t_0) = \left[r_{\text{dif}}(t_0)s \{ \rho P_1 + (1 - \rho)P_2 \} - (1 - \rho)P_1 \right] \bar{I}_{\text{dif}}(t_0) A_{\text{dif}}^2(t_0) - \rho r_{\text{dif}}(t_0)s \bar{I}_{\text{dif}}(t_0) C_{\text{dif}}^2(t_0). \quad (59)$$

ここで、 $r_{\text{dif}}(t)$ は確率分布関数 $I_{\text{dif}}(t)$ のハザード率であり、 $r_{\text{dif}}(t) = [dI_{\text{dif}}(t)/dt] / \bar{I}_{\text{dif}}(t)$ で定義される。いま、式 (56) から $r_{\text{dif}}(t)$ に関して以下の補題が成り立つ。

補題 7 [6]: 以下の諸量を定義する。

$$t_m \equiv -\frac{3b_1}{2a_1^2} + \frac{1}{a_1} \sqrt{1 + \frac{9b_1^2}{4a_1^2}}, \quad (60)$$

$$t_n \equiv \frac{2}{3b_1}. \quad (61)$$

このとき方程式

$$r_{\text{dif}}(t) = \frac{a_1^2}{2b_1} + \frac{3}{2t} - \frac{1}{2b_1 t^2} \quad (62)$$

を満足する $t_e \in (t_m, t_n]$ が存在すれば, $r_{\text{dif}}(t)$ は $t = t_e$ で最大値をとる単峰関数となる. \square

以上の準備の下で, 電力有効性 $W_{\text{dif}}^2(t_0)$ を最大にする最適停止時間 t_0^* に関する次の結果が得られる.

定理 8: 仮定 $\rho/[\rho P_1 + (1-\rho)P_2] < W_{\text{dif}}^2(t_0)$ の下で $t_e \in (t_m, t_n]$ が存在すると仮定する.

1) $q_{\text{dif}}^2(t_e) > 0$ かつ $q_{\text{dif}}^2(\infty) < 0$ のとき,

i) $q_{\text{dif}}^2(0) > 0$ ならば, 電力有効性を最大にする有限で唯一の最適停止時間 $t_0^* \in (t_e, \infty)$ が存在し, $q_{\text{dif}}^2(t_0^*) = 0$ を満足する. そのときの最大電力有効性は,

$$W_{\text{dif}}^2(t_0^*) = \rho r_{\text{dif}}(t_0^*)s / \left[\{\rho P_1 + (1-\rho)P_2\} \times r_{\text{dif}}(t_0^*)s - (1-\rho)P_1 \right] \quad (63)$$

となる.

ii) $q_{\text{dif}}^2(0) \leq 0$ ならば $q_{\text{dif}}^2(\bar{t}_0) = 0$ を満足する $\bar{t}_0 \in (t_e, \infty)$ が存在し, 不等式

$$r_{\text{dif}}(\bar{t}_0) < \frac{P_1[\rho + \lambda\tau + \lambda\rho s]}{P_2 s[\rho + \lambda\tau]} \quad (64)$$

を満たすとき, 最適停止時間は $t_0^* = \bar{t}_0$ となり, 最小期待電力は式 (63) で与えられる. そうでない場合は $t_0^* = 0$ となる.

2) $q_{\text{dif}}^2(t_e) \leq 0$ または $q_{\text{dif}}^2(\infty) \geq 0$ のとき,

$$\frac{P_2}{P_1} < \frac{1}{\lambda s} + \frac{\rho}{\rho + \lambda\tau} \quad (65)$$

ならば $t_0^* = 0$ となり, そうでなければ $t_0^* \rightarrow \infty$ となる. \square

4. 今後の課題

本稿では, コンピュータシステムの自動スリープ機能に対して, アクセスの処理中に発生するアクセス要求がキャンセルされるモデルと待ち行列を形成する2種類のモデルを考察した. アクセス要求の発生がポアソン過程に従う場合では, 最適停止時間を解析的に導出した. また, アクセス要求の発生が一

般の再生過程に従う場合に対しては, 電力有効性を近似的に導出した. 特に, 待ち行列を形成するモデルにおいて, アクセス要求の発生が再生過程に従う場合に対しては, 拡散近似を適用することで近似的な最適解を導出した. 今後は, 各近似手法の精度の評価を行う予定である.

参考文献

- [1] H. Sandoh, H. Hirakoshi and H. Kawai, "An optimal time to sleep for an auto-sleep system", *Computeres & Operations Research*, Vol. 23, pp. 221-227 (1996).
- [2] 土肥, 海生, 尾崎, "ノート型パーソナルコンピュータの省電力方策に対するノンパラメトリックアプローチ", *信学論*, Vol. J78-A, pp. 1157-1165 (1995).
- [3] M. Miyazawa, "A generalizad Pollaczek-Khintchine formula for the $GI/GI/1/K$ queue and its application to approximation", *Stochastic Models*, Vol. 3, pp. 53-65 (1989).
- [4] T. Kimura, "Optimal control of an $M/G/1$ queuing system with removable server via diffusion approximation", *European Journal of Operational Research*, Vol. 8, pp. 390-398 (1981).
- [5] H. Okamura, T. Dohi and S. Osaki, "Optimal timing strategies for controlled $M/G/1$ queuing system via diffusion approximation approach", in submission.
- [6] R. S. Chhikara and J. L. Folks, "The inverse Gaussian distribution as a lifetime model", *Technometrics*, Vol. 19, pp. 461-468 (1977).