

コンピュータシステムの自動スリープ機能による省電力効果 II ——待ち行列モデル

岡村 寛之^{†,☆} 土肥 正[†] 尾崎 俊治[†]

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。本論文では、コンピュータへのアクセスがない状況において、ユーザがいつハードディスクをスリープ状態に切り替えるべきかについての問題を取りあげ、定常状態における単位時間あたりの期待消費電力を最小にする最適スリープ時刻を求めるための確率モデルを提案する。特に、ユーザによるコンピュータへのアクセス要求が到着した後、その処理期間中に到着する他のアクセス要求がバッファ内で待ち行列を形成するようなモデルについて考察を行う。特別な場合として、アクセス要求が同次ポアソン過程に従って発生し、バッファ内のアクセス要求数が $M/G/1$ 待ち行列を形成する場合、期待消費電力を最小にする最適スリープ時刻が解析的に導出される。より一般的な到着パターンに対しては、3種類の近似解が提供される。

On the Effect of Power Saving by Auto-sleep Function of a Computer System II—Queueing Model

HIROYUKI OKAMURA,^{†,☆} TADASHI DOHI[†] and SHUNJI OSAKI[†]

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. This paper addresses a problem of how to determine the optimal sleep timing when the computer user should turn the hard disk or the display off in order to save the electric power after the computer has not been accessed. We propose a stochastic model to obtain the optimal sleep timing strategy which minimizes the expected power consumed per unit time in the steady-state, when the access requirements arrive at a computer according to a renewal process. In this model, if the first transaction is accessed, the other transactions arrived later are cumulated in a buffer until completing the process of the first one. We derive the optimal auto-sleep strategy analytically when the access requirements arrive in accordance with a homogeneous Poisson process and the buffer behaviour follows an $M/G/1$ queueing process, and provide three approximated solutions for more general cases.

1. はじめに

近年、コンピュータを構成する部品や周辺機器の高速化・高機能化にともなう性能の向上により、我々の周りにおけるコンピュータの利用機会は急激に広まっている。このような状況に対して、コンピュータの信頼性や操作性に加え、コンピュータのパワーマネジメントと呼ばれる省電力技術の重要性が注目されるようになってきている。パワーマネジメントの必要性は、米国の環境保護庁省エネルギー奨励策 (Energy Star

Computers Program: EPS) が発表されて以来世界的な関心を集めており、わが国においてもすでに ESP 基準に適合したワークステーションが販売されるに至っている。また、コンピュータの省電力化の需要が最も顕著な例として、ノートブック PC があげられる。ノートブック PC の重要な特徴である高い携帯性を実現するために、バッテリー駆動時間の延長だけでなく低電圧かつ高速クロックで駆動する回路を最小化するデバイスの開発など、多くの未解決な問題が山積みされているのが現状である^{1),2)}。

本論文では、ハードウェア本体やアプリケーションプログラムをアイドル状態から低消費電力状態へ移行させるための自動スリープ機能 (auto-sleep function) について考察を行う。このような自動スリープ機能は、最近のほとんどのデスクトップ PC ならびにノート

[†] 広島大学工学部第二類 (電気系)

Faculty of Engineering, Cluster II (Electrical and Industrial Engineering), Hiroshima University

[☆] 現在、株式会社 CSK

Presently with CSK Corporation

ブック PC に標準装備されており、コンピュータ設計における基盤技術として定着しつつある。前報³⁾において著者らは、コンピュータへのアクセス要求がない状況において、省電力の観点からユーザがいつハードディスクをスリープ状態に切り替えるべきかについての問題を定式化し、定常状態における単位時間あたりの期待消費電力を最小にする最適スリープ時刻を求めるための確率モデルを提案した。そこで考えられたモデルは、ユーザによるコンピュータへのアクセス要求が到着した後、その処理期間中に到着する他のアクセス要求は自動的にキャンセルされると仮定しており、マルチユースを想定したデスクトップ型コンピュータの使用形態をモデル化したものであった。

しかしながら、コンピュータネットワークとオンラインランザクシオン処理が普及している現在、あらゆるシステムにおいて上述のような処理がつねに行われているとは考えにくい。たとえば、マルチタスクの OS を搭載しているコンピュータシステムに接続されているハードディスクやネットワークプリンタなどの自動スリープ機能を設計する場合、到着したジョブはバッファ内に蓄積され、到着順に逐次処理されなければならない。このような背景に鑑み、本論文ではアクセス要求がバッファ内で待ち行列を形成するようなシステムについて考察を行い、定常状態における単位時間あたりの期待消費電力を最小にする最適スリープ時刻を求めるための異なる確率モデルを提案する。

本論文の構成は次のとおりである。まず次章においてモデルの概要を述べた後、3章においてコンピュータへのアクセス要求の発生回数が再生過程に従って生起し、バッファ内のアクセス要求数が $GI/G/1$ 待ち行列を形成する場合の定常状態における単位時間あたりの期待消費電力を定式化する。この場合、最適自動スリープ時刻を解析的に求めることはきわめて困難となる。そこで4章において、アクセス要求の発生回数が同次ポアソン過程に従って生起し、バッファ内のアクセス要求数が $M/G/1$ 待ち行列を形成する特別な場合に着目し、期待消費電力を最小にする最適自動スリープ時刻を解析的に導出する。続いて5章では、アクセス要求の一般的な到着パターンに対して2種類の近似解法を提案する。さらに6章では、待ち行列過程を精度良く近似するために、拡散近似の手法を適用することを考える。最終的に7章において、最適自動スリープ時刻を数値的に求め、提案した近似解法の精度を比較するための数値実験を行う。

2. コンピュータシステムの自動スリープ機能

コンピュータシステムにおいて省電力化を効率的に実現するための1つの手段である自動スリープ機能について述べる。最初に、モデル化を行うために必要な記号の定義を行う。

- s : コンピュータへのアクセス (call) が行われた後、コンピュータを起動するために必要な時間 (正定数)
- τ : アクセス要求後、実際に処理を開始するまでに要する準備時間 (正定数)
- X_k : k ($= 1, 2, \dots$) 番目のアクセス要求発生時間間隔 (非負値確率変数)
- S_k : k 番目のアクセス要求 (ランザクシオン) に対する処理時間 (非負値確率変数)
- t_0 : スリープ開始時間 ($0 \leq t_0 < \infty$)
- P_1 : コンピュータの稼働状態において発生する単位時間あたりの消費電力
- P_2 : コンピュータの起動状態において発生する単位時間あたりの消費電力 ($P_2 > P_1 > 0$)。

また、Sandoh⁴⁾に従って、本論文で対象とする自動スリープ機能を有するシステムは、以下に示す4つの状態を推移するものと仮定する (図1参照)。

スリープモード (sleep): 電源 OFF とは異なるシステムの休止状態であり、アクセス要求が発生すると起動状態へただちに移行する。スリープモードにおいて発生する消費電力は P_1 や P_2 と比較して相対的に非常に小さいことが知られており、本論文ではスリープモードにおける消費電力は0であると仮定する。

起動状態 (warm-up): システムが起動中の状態であり、ある一定期間 ($0, s$] 経過すると待機状態へ移行する。この状態での消費電力は通常の消費電力と比較して非常に大きく、単位時間あたり P_2 の起動電力が消費される。

待機状態 (idle): アクセス要求に対して待機している状態であり、アクセス要求が発生するとただちに稼働状態へ移行する。ここで、ある一定期間

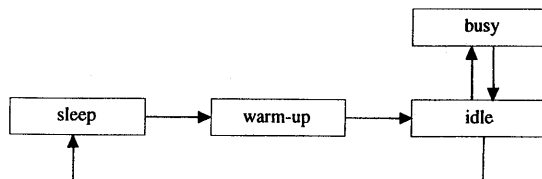


図1 自動スリープ機能の概念図

Fig. 1 Configuration of an auto-sleep system.

$(0, t_0]$ 中アクセス要求が発生しない場合、ただちにスリープモードへ移行する。待機状態では単位時間あたり P_1 の通常電力が消費される。

稼働状態 (busy): システムが実際にアクセスに対する処理を行っている状態であり、アクセスの処理を開始するまでに要する準備時間 $(0, \tau]$ を経過した後、処理を開始する。トランザクションの処理が完了すると待機状態へ移行する。稼働状態では単位時間あたり P_1 の通常電力が消費される。

いま、コンピュータに対するアクセス要求は確率的に生起するものとし、 $k-1$ 番目のアクセス要求の発生から k 番目のアクセス要求の発生までの時間間隔 X_k ($k = 1, 2, \dots$) は互いに独立で同一な確率変数列とする。すなわち、アクセス要求回数 $\{N(t), t \geq 0\}$ は一般の再生過程⁵⁾に従うものと仮定する。アクセス要求発生時間間隔 X_k の確率分布関数を $F(t) = \Pr\{X_k \leq t\}$ 、平均と分散を、それぞれ $E[X_k] = 1/\lambda$ ならびに $\text{Var}[X_k] = \sigma_a^2$ とする。アクセス要求に対する処理は先着順 (first-in first-out: FIFO) に行われ、コンピュータはバッファ内に蓄積している要求がすべてなくなるまで処理を継続する。ここで、単一のアクセス要求に対する処理時間 S_k ($k = 1, 2, \dots$) は互いに独立で同一な確率変数列とし、 S_k の確率分布関数を $G(t) = \Pr\{S_k \leq t\}$ 、平均と分散を、それぞれ $E[S_k] = 1/\mu$ ならびに $\text{Var}[S_k] = \sigma_s^2$ と表記する。以降では簡単のため、 X_k および S_k の添字 k を省略し、 X および S のように略記する。

バッファ内のアクセス要求がすべて処理された時点から時間 t_0 経過してもアクセス要求がまったくない場合、コンピュータはスリープモードへと移行する。コンピュータがスリープモードにあるときにアクセス要求が発生すると、起動時間 s を経過してコンピュータは稼働状態となる。コンピュータシステムの起動開始時点から次の起動開始時点までの期間を1サイクルとすれば、各々の起動開始時点は再生点⁵⁾となる。

3. GI/G/1 待ち行列によるモデル化

前章で述べたように、アクセス要求の発生時間間隔が独立で同一な確率分布関数 $F(t)$ に従い、各々の処理時間が一般の確率分布関数 $G(t)$ に従うとき、コンピュータのバッファ内に蓄積されるアクセス要求数が形成する待ち行列過程は $GI/G/1/FIFO$ 待ち行列 (以下では、特に断らない限り $GI/G/1$ と略記) となる。ここで、通常の $GI/G/1$ 待ち行列ではサーバが空のときに客が到着 (アクセス要求が発生) すると即座にサービス (アクセスに対する処理) を行うのに対

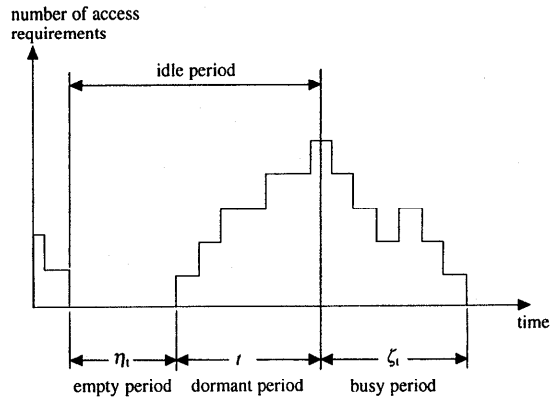


図2 バケーションをともなう待ち行列システム
Fig. 2 Queueing system with vacation.

して、本論文で取り扱うモデルでは、サーバが空のときにアクセス要求が発生した場合、時間 $s + \tau$ もしくは τ を経過して実際にアクセス要求に対する処理を行う。このような時間遅れを持つ待ち行列はバケーションをともなう待ち行列 (vacation queue)⁶⁾ とよばれ、応用上きわめて重要な待ち行列過程である。

バケーションをともなう待ち行列において、サーバの空き期間 (empty period) とサーバの稼働時間 (busy period) はサーバの余暇期間 (dormant period) に依存する (図2 参照)。いま、一般性を失うことなく、 $GI/G/1$ 待ち行列のトラフィック強度 (利用率) に関して $\rho = \lambda/\mu < 1$ が仮定できるものとするれば、以下の関係式が成立することが知られている⁶⁾。

$$E[\text{busy period}] = \rho \left\{ E[\text{dormant period}] + E[\text{busy period}] + E[\text{empty period}] \right\}. \quad (1)$$

式(1)を利用して、自動スリープシステムの定常状態における単位時間あたりの期待消費電力を表現するために、次のような確率変数を定義する。

η_t : サーバの空き期間

$$I(x|t) = \Pr\{\eta_t \leq x\}$$

ζ_t : コンピュータがアクセス要求に対して処理を行う期間。

ここで、確率変数 η_t と ζ_t の添字 t は、バケーションをともなう待ち行列におけるサーバの余暇期間を示している。これは、自動スリープシステムにおいてバッファ内にアクセス要求が発生しているにもかかわらず、アクセス処理を行わない期間、すなわちアクセス処理に対する準備時間 τ もしくは起動開始から準備完了までの時間 $s + \tau$ に対応する。したがって、式(1)より以下の関係式を得る。

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

いま、稼働状態でアクセス要求が発生した直後にスリープモードが終了するまでの期待時間を $T_{act}(t_0)$ 、1 サイクルの期待時間を $T(t_0)$ とすれば、次のような関係式が成り立つ。

$$\begin{aligned} T(t_0) = & \int_0^{t_0} \left\{ s + \tau + E[\zeta_{s+\tau}] + t + T_{act}(t_0) \right\} \\ & \times dI(t | s + \tau) \\ & + \int_{t_0}^{\infty} \left\{ s + \tau + E[\zeta_{s+\tau}] + t \right\} \\ & \times dI(t | s + \tau), \end{aligned} \quad (3)$$

$$\begin{aligned} T_{act}(t_0) = & \int_0^{t_0} \left\{ \tau + E[\zeta_\tau] + t + T_{act}(t_0) \right\} \\ & \times dI(t | \tau) \\ & + \int_{t_0}^{\infty} \left\{ \tau + E[\zeta_\tau] + t \right\} dI(t | \tau). \end{aligned} \quad (4)$$

式(3)と式(4)から、1 サイクルの期待時間は

$$\begin{aligned} T(t_0) = & s + \tau + E[\zeta_{s+\tau}] + E[\eta_{s+\tau}] \\ & + \left\{ \tau + E[\zeta_\tau] + E[\eta_\tau] \right\} \frac{I(t_0 | s + \tau)}{\bar{I}(t_0 | \tau)}. \end{aligned} \quad (5)$$

ここで、 $\bar{I}(\cdot) = 1 - I(\cdot)$ は余関数を表す。さらに式(2)を用いて、 $E[\zeta_{s+\tau}]$ および $E[\zeta_\tau]$ を消去すると、1 サイクルの期待時間は

$$\begin{aligned} T(t_0) = & \frac{1}{1 - \rho} \left\{ s + \tau + E[\eta_{s+\tau}] \right. \\ & \left. + \left(\tau + E[\eta_\tau] \right) \frac{I(t_0 | s + \tau)}{\bar{I}(t_0 | \tau)} \right\} \end{aligned} \quad (6)$$

となる。図3にモデルの概念図を示す。図3において示される1サイクル中のサブサイクルが発生する回数(確率変数)を N とすれば、 N は1回目の成功する確率が $\Pr\{\eta_{s+\tau} \leq t_0\}$ 、2回目以降で成功する確率が $\Pr\{\eta_\tau \leq t_0\}$ の幾何分布に従うため、 N の期待値は

$$E[N] = \frac{\Pr\{\eta_{s+\tau} \leq t_0\}}{\Pr\{\eta_\tau > t_0\}} = \frac{I(t_0 | s + \tau)}{\bar{I}(t_0 | \tau)} \quad (7)$$

となる。これより、1サイクルの期待時間は、 N を用いて、

$$\begin{aligned} T(t_0) = & \frac{1}{1 - \rho} \left\{ s + \tau + E[\eta_{s+\tau}] + E[N] \right. \\ & \left. \times \left(\tau + E[\eta_\tau] \right) \right\} \end{aligned} \quad (8)$$

となる。

同様に、期間 $T_{act}(t_0)$ 中に生じる総期待消費電力を $C_{act}(t_0)$ 、1 サイクル中に生じる総期待消費電力

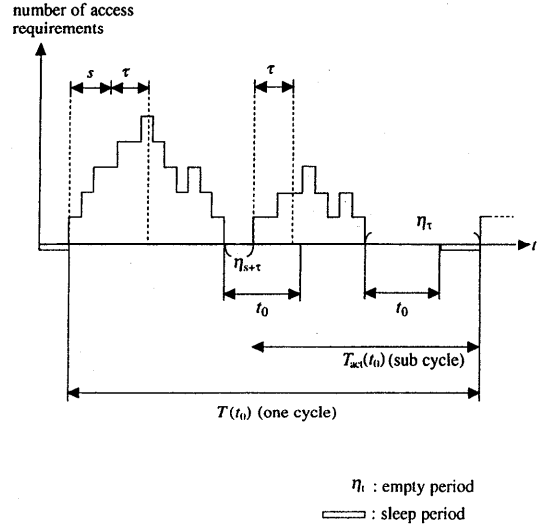


図3 確率システムの挙動例

Fig. 3 Possible realization of the stochastic system.

を $C(t_0)$ とすれば、

$$\begin{aligned} C(t_0) = & \int_0^{t_0} \left\{ P_2 s + P_1 (\tau + E[\zeta_{s+\tau}] + t) \right. \\ & \left. + C_{act}(t_0) \right\} dI(t | s + \tau) \\ & + \int_{t_0}^{\infty} \left\{ P_2 s + P_1 (\tau + E[\zeta_{s+\tau}] + t_0) \right\} \\ & \times dI(t | s + \tau), \end{aligned} \quad (9)$$

$$\begin{aligned} C_{act}(t_0) = & \int_0^{t_0} \left\{ P_1 (\tau + E[\zeta_\tau] + t) + C_{act}(t_0) \right\} \\ & \times dI(t | \tau) + \int_{t_0}^{\infty} P_1 (\tau + E[\zeta_\tau] + t_0) \\ & \times dI(t | \tau) \end{aligned} \quad (10)$$

を得る。

式(9)と(10)から、1サイクルにおける総期待消費電力は

$$\begin{aligned} C(t_0) = & P_2 s + P_1 \left\{ \tau + E[\zeta_{s+\tau}] + E[\eta_{s+\tau} \wedge t_0] \right\} \\ & + P_1 \left\{ \tau + E[\zeta_\tau] + E[\eta_\tau \wedge t_0] \right\} \\ & \times \frac{I(t_0 | s + \tau)}{\bar{I}(t_0 | \tau)}. \end{aligned} \quad (11)$$

ここで、

$$\begin{aligned} E[\eta_t \wedge t_0] = & E[\min(\eta_t, t_0)] \\ = & \int_0^{t_0} u dI(u | t) + t_0 \bar{I}(t_0 | t) \end{aligned} \quad (12)$$

である。式(2)から、 $E[\zeta_{s+\tau}]$ および $E[\zeta_\tau]$ を消去すると、1サイクルにおける総期待消費電力は、 N を用いて、

$$\begin{aligned}
C(t_0) = & \left\{ \frac{\rho}{1-\rho} P_1 + P_2 \right\} s + \frac{P_1 \tau}{1-\rho} + P_1 \\
& \times \left\{ \frac{\rho}{1-\rho} E[\eta_{s+\tau}] + E[\eta_{s+\tau} \wedge t_0] \right\} \\
& + E[N] \left\{ \frac{P_1 \tau}{1-\rho} + P_1 \left(\frac{\rho}{1-\rho} E[\eta_r] \right. \right. \\
& \left. \left. + E[\eta_r \wedge t_0] \right) \right\} \quad (13)
\end{aligned}$$

となる。これより、定常状態における単位時間あたりの期待消費電力は、よく知られた再生定理⁵⁾を用いて

$$\begin{aligned}
V(t_0) = & \lim_{t \rightarrow \infty} \frac{E[(0, t] \text{ における総消費電力}]}{t} \\
= & C(t_0)/T(t_0) \quad (14)
\end{aligned}$$

となる。これより、問題は

$$\min_{0 \leq t_0 < \infty} V(t_0) \quad (15)$$

によって定式化される。式(8)と(13)から、定常状態における単位時間あたりの期待消費電力を陽に表現するためには、確率変数 η_t に関する性質を特定しなければならない。次章では特別な場合として、アクセス要求の発生回数が同次ポアソン過程に従う場合、すなわちバッファ内に蓄積されるアクセス要求の発生回数が $M/G/1$ 待ち行列過程に従う場合について解析を行う。これは、アクセス要求が時間に関して一定の到着率に従って生起する場合に対応しており、最も基本的であるがきわめて現実的な状況をモデル化したものである。

4. $M/G/1$ 待ち行列の場合

ここでは、 $GI/G/1$ 待ち行列の特別な場合として、アクセス要求の発生回数がパラメータ $\lambda (> 0)$ のポアソン過程 $\{N(t), t \geq 0\}$ に従う場合を考える。すなわち、

$$F(t) = 1 - \exp(-\lambda t). \quad (16)$$

このとき、 $M/G/1$ 待ち行列におけるサーバの空き期間分布は

$$I(x | t) = 1 - \exp(-\lambda x) \quad (17)$$

の指数分布となるので、

$$E[\eta_t] = \frac{1}{\lambda}, \quad (18)$$

$$E[\eta_t \wedge t_0] = \frac{1 - \exp(-\lambda t_0)}{\lambda} \quad (19)$$

を得る。式(18)を式(6)に代入することによって、1サイクルの期待時間は

$$T_{M/G/1}(t_0) = \frac{\tau + 1/\lambda}{(1-\rho) \exp(-\lambda t_0)} + \frac{s}{1-\rho}. \quad (20)$$

また、式(19)を式(13)に代入することによって、1

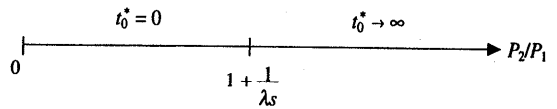


図4 ポアソン到着の場合に対する最適スリープ戦略

Fig. 4 Optimal sleeping strategy for the Poisson arrival case.

サイクル中に生じる総期待消費電力は

$$\begin{aligned}
C_{M/G/1}(t_0) = & \frac{P_1}{(1-\rho) \exp(-\lambda t_0)} \\
& \times \left\{ \tau + \frac{1 - (1-\rho) \exp(-\lambda t_0)}{\lambda} \right\} \\
& + \frac{\rho P_1 + (1-\rho) P_2}{1-\rho} s \quad (21)
\end{aligned}$$

となる。これより、定常状態における単位時間あたりの期待消費電力は

$$V_{M/G/1}(t_0) = C_{M/G/1}(t_0)/T_{M/G/1}(t_0) \quad (22)$$

となる。このとき、定常状態における単位時間あたりの期待消費電力 $V_{M/G/1}(t_0)$ を最小にする最適スリープ時間 t_0^* に関して次の結果が得られる。

定理 4.1: 条件 $\rho < 1$ を仮定する。もし $P_2/P_1 < 1 + 1/(\lambda s)$ ならば、定常状態における単位時間あたりの期待消費電力 $V_{M/G/1}(t_0)$ を最小にする最適スリープ時間は $t_0^* = 0$ 、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、そのときの最小期待消費電力は

$$V_{M/G/1}(0) = \frac{P_1(\tau + 1/\mu) + (\rho P_1 + (1-\rho) P_2)s}{s + \tau + 1/\lambda} \quad (23)$$

となる。逆に $P_2/P_1 \geq 1 + 1/(\lambda s)$ ならば $t_0^* \rightarrow \infty$ 、すなわちまったくスリープモードに入らないことが最適となり、そのときの最小期待消費電力は

$$V_{M/G/1}(\infty) = P_1 \quad (24)$$

となる。

定理 4.1 の証明は付録に与えられる。上述の結果より、コンピュータへのアクセス要求の発生頻度が時間に関して一定である場合、トランザクションの処理終了後ただちにスリープモードに移行するか、まったくスリープを行わないことが最適となる。図 4 は、バッファの時間的挙動が $M/G/1$ 待ち行列過程に従う場合の最適スリープ時間とパラメータ比 P_2/P_1 の関係を示したものである。結果は前報³⁾におけるポアソン到着の場合と同じであり、最適スリープ時間はアクセス要求到着率 λ とアクセス準備時間 s に従属することが分かる。

5. 空き期間分布の近似

アクセス要求発生時間間隔分布 $F(t)$ が一般分布であるときの最適スリープ時間を導出するためには、サーバの空き期間 η_t の確率分布関数 $I(x|t)$ を特定しなければならない。しかしながら一般的に、 $GI/G/1$ 待ち行列の空き期間分布 $I(x|t)$ を求めることは非常に困難である。そこで、 $I(x|t)$ を何らかの関数によって近似することにより、式 (14) で表現された定常状態における単位時間あたりの期待消費電力 $V(t_0)$ の近似式を導出する。以下では、Miyazawa⁷⁾ によって提案された2つの近似手法を適用することを試みる。

5.1 均衡分布による近似 (近似1)

サーバの空き期間分布をアクセス要求時間間隔の均衡分布を用いて近似する。すなわち、

$$I(x|t) \approx I_{APP1}(x) \\ = \lambda \int_0^x \bar{F}(u) du. \quad (25)$$

これより、

$$E[\eta_t] \approx E[\eta_{APP1}] = \frac{\lambda^2 \sigma_a^2 + 1}{2\lambda}, \quad (26)$$

$$E[\eta_t \wedge t_0] \approx E[\eta_{APP1} \wedge t_0] \\ = \lambda \int_0^{t_0} \int_x^\infty \bar{F}(u) du dx \quad (27)$$

を得る。関数 $I_{APP1}(x)$ が t に依存しないことに注意して、式 (6) の1サイクルの期待時間と式 (13) の1サイクル中に生じる総期待消費電力に式 (26) と (27) の近似式を代入すると、1サイクルの期待時間の近似式は

$$T(t_0) \approx T_{APP1}(t_0) \\ = \frac{1}{(1-\rho)\bar{I}_{APP1}(t_0)} \left\{ \tau + E[\eta_{APP1}] \right\} \\ + \frac{s}{1-\rho} \quad (28)$$

となり、1サイクル中に生じる総期待消費電力の近似式は

$$C(t_0) \approx C_{APP1}(t_0) \\ = \frac{P_1}{\bar{I}_{APP1}(t_0)} \left\{ \frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\eta_{APP1}] \right. \\ \left. + E[\eta_{APP1} \wedge t_0] \right\} + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s \quad (29)$$

となる。これらを用いて、定常状態における単位時間あたりの期待消費電力は

$$V(t_0) \approx V_{APP1}(t_0) = C_{APP1}(t_0)/T_{APP1}(t_0) \quad (30)$$

となる。

いま、次のような非線形関数

$$q_{APP1}(t_0) \equiv \left\{ (1-\rho)P_1 - r_{APP1}(t_0)s(\rho P_1 \right. \\ \left. + (1-\rho)P_2) \right\} \bar{I}_{APP1}(t_0) T_{APP1}(t_0) \\ + r_{APP1}(t_0)s \bar{I}_{APP1}(t_0) C_{APP1}(t_0) \quad (31)$$

を定義する。ここで、 $r_{APP1}(t)$ は $I_{APP1}(t)$ のハザード率であり、

$$r_{APP1}(t) \equiv \frac{dI_{APP1}(t)/dt}{\bar{I}_{APP1}(t)} = \frac{\bar{F}(t)}{\int_t^\infty \bar{F}(u) du} \quad (32)$$

である。

以上より、 $V_{APP1}(t_0)$ を最小にする最適スリープ時間 t_0^* に関する定理が得られる。

定理5.1 (近似1): 仮定 $\rho P_1 + (1-\rho)P_2 > V_{APP1}(t_0)$ の下で、(1) アクセス要求時間間隔の確率分布関数 $F(t)$ が狭義 DHR (Decreasing Hazard Rate) とする。

(i) $q_{APP1}(0) < 0$ かつ $q_{APP1}(\infty) > 0$ ならば、定常状態における単位時間あたりの期待消費電力を最小にする有限で唯一の最適スリープ時間の近似値 $t_0^* \in (0, \infty)$ が存在し、非線形方程式 $q_{APP1}(t_0^*) = 0$ を満足する。そのときの最小期待消費電力は、

$$V_{APP1}(t_0^*) = \rho P_1 + (1-\rho)P_2 - \frac{(1-\rho)P_1}{r_{APP1}(t_0^*)s} \quad (33)$$

となる。

(ii) $q_{APP1}(0) \geq 0$ ならば最適スリープ時間は $t_0^* = 0$ となり、最小期待消費電力は

$$V_{APP1}(0) = \frac{P_1(\tau + \rho E[\eta_{APP1}])}{\tau + E[\eta_{APP1}] + s} \\ + \frac{\{\rho P_1 + (1-\rho)P_2\}s}{\tau + E[\eta_{APP1}] + s} \quad (34)$$

となる。

(iii) $q_{APP1}(\infty) \leq 0$ ならば最適スリープ時間は $t_0^* \rightarrow \infty$ となり、最小期待消費電力は

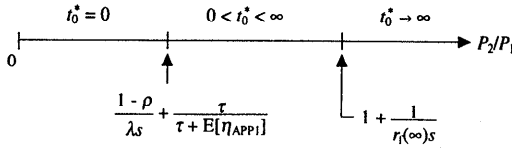
$$V_{APP1}(\infty) = P_1 \quad (35)$$

となる。

(2) アクセス要求時間間隔の確率分布関数 $F(t)$ が IHR (Increasing Hazard Rate) とする。もし、 $P_2/P_1 < 1 + E[\eta_{APP1}]/s$ ならば $t_0^* = 0$ となり、そうでなければ $t_0^* \rightarrow \infty$ となる。

証明は定理4.1と同様であるので省略する。定理5.1で仮定された条件 $\rho P_1 + (1-\rho)P_2 > V_{APP1}(t_0)$ は前

(I) DHR case



(II) IHR case

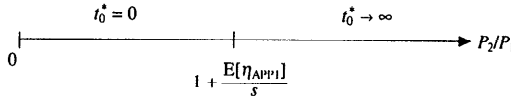


図5 近似1に基づいた最適スリープ戦略
Fig. 5 Optimal sleeping strategy based on approximation 1.

報³⁾の定理 5.1 で仮定されたもの ($P_2 > V_{APP1}(t_0)$) よりも制約的であることに注目すべきである。すなわち、トラフィック強度 ρ が十分小さいならば、設定された仮定はきわめて現実的であるといえる。しかしながら、アクセス要求到着率が処理率にかなり近い値をとる状況においては、もはや解析的に最適スリープ時刻の存在を保証することが困難となる。図 5 は定理 5.1 の内容を視覚的に図示したものである。ここで、図中の $r_f(t) = (dF(t)/dt)/\bar{F}(t)$ はアクセス要求時間間隔分布のハザード率である。これより、仮定 $\rho P_1 + (1 - \rho)P_2 > V_{APP1}(t_0)$ の下で、最適スリープ時間は電力比 P_2/P_1 に対して単調な性質を持つことが分かる。

5.2 指数分布による近似 (近似 2)

次に、空き期間分布に対して次のような指数分布を利用した近似を適用する。

$$I(x | t) \approx I_{APP2}(x) = 1 - \frac{\int_0^\infty \exp(-\mu u) \bar{F}(x+u) du}{\int_0^\infty \exp(-\mu u) \bar{F}(u) du} \quad (36)$$

この近似を適用することにより、

$$E[\eta_t] \approx E[\eta_{APP2}] = \frac{1}{\lambda \mu \int_0^\infty \exp(-\mu u) \bar{F}(u) du} - \frac{1}{\mu} \quad (37)$$

$$E[\eta_t \wedge t_0] \approx E[\eta_{APP2} \wedge t_0] = \frac{\int_0^{t_0} \int_0^\infty \exp(-\mu u) \bar{F}(x+u) du dx}{\int_0^\infty \exp(-\mu u) \bar{F}(u) du} \quad (38)$$

$I_{APP2}(x)$ が t に依存しないことに注意して、式 (6)

と式 (13) に式 (36)~(38) の近似式を代入すると、1 サイクルの期待時間の近似式は

$$T(t_0) \approx T_{APP2}(t_0) = \frac{1}{(1-\rho)\bar{I}_{APP2}(t_0)} (\tau + E[\eta_{APP2}]) + \frac{s}{1-\rho} \quad (39)$$

となり、1 サイクル中に生じる総期待消費電力の近似式は

$$C(t_0) \approx C_{APP2}(t_0) = \frac{P_1}{\bar{I}_{APP2}(t_0)} \left\{ \frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\eta_{APP2}] + E[\eta_{APP2} \wedge t_0] \right\} + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s \quad (40)$$

となる。これらを用いて、定常状態における単位時間あたりの期待消費電力は

$$V(t_0) \approx V_{APP2}(t_0) = C_{APP2}(t_0)/T_{APP2}(t_0) \quad (41)$$

となる。

式 (31) と同様に、非線形関数

$$q_{APP2}(t_0) \equiv \left\{ (1-\rho)P_1 - r_{APP2}(t_0)s\{\rho P_1 + (1-\rho)P_2\} \right\} \bar{I}_{APP2}(t_0) T_{APP2}(t_0) + r_{APP2}(t_0)s \bar{I}_{APP2}(t_0) C_{APP2}(t_0) \quad (42)$$

を定義する。ここで、 $r_{APP2}(t)$ は $I_{APP2}(t)$ のハザード率であり、

$$r_{APP2}(t) \equiv \frac{dI_{APP2}(t)/dt}{\bar{I}_{APP2}(t)} = \frac{\exp(-\mu t) \bar{F}(t)}{\int_t^\infty \exp(-\mu u) \bar{F}(u) du} - \mu \quad (43)$$

である。

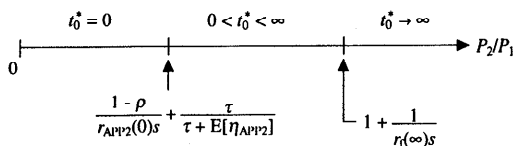
以上より、 $V_{APP2}(t_0)$ を最小にする最適スリープ時間 t_0^* に関する以下の定理を得る。

定理 5.2 (近似 2) : $\rho P_1 + (1 - \rho)P_2 > V_{APP2}(t_0)$ の下で、(1) アクセス要求時間間隔の確率分布関数 $F(t)$ が狭義 DHR とする。

(i) $q_{APP2}(0) < 0$ かつ $q_{APP2}(\infty) > 0$ ならば、定常状態における単位時間あたりの期待消費電力を最小にする有限で唯一の最適スリープ時間の近似値 $t_0^* \in (0, \infty)$ が存在し、非線形方程式 $q_{APP2}(t_0^*) = 0$ を満足する。そのときの最小期待消費電力は、

$$V_{APP2}(t_0^*) = \rho P_1 + (1 - \rho)P_2$$

(I) DHR case



(II) IHR case

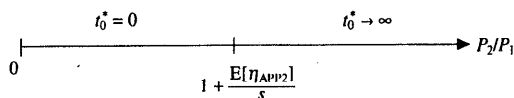


図6 近似2に基づいた最適スリープ戦略

Fig. 6 Optimal sleeping strategy based on approximation 2.

$$-\frac{(1-\rho)P_1}{r_{APP2}(t_0^*)s} \quad (44)$$

となる。

- (ii) $q_{APP2}(0) \geq 0$ ならば最適スリープ時間は $t_0^* = 0$ となり、最小期待消費電力は

$$V_{APP2}(0) = \frac{P_1(\tau + \rho E[\eta_{APP2}])}{\tau + E[\eta_{APP2}] + s} + \frac{\{\rho P_1 + (1-\rho)P_2\}s}{\tau + E[\eta_{APP2}] + s} \quad (45)$$

となる。

- (iii) $q_{APP2}(\infty) \leq 0$ ならば最適スリープ時間は $t_0^* \rightarrow \infty$ となり、最小期待消費電力は

$$V_{APP2}(\infty) = P_1 \quad (46)$$

となる。

- (2) アクセス要求時間間隔の確率分布関数 $F(t)$ が IHR とする。

- (i) $P_2/P_1 < 1 + E[\eta_{APP2}]/s$ ならば $t_0^* = 0$,
(ii) $P_2/P_1 \geq 1 + E[\eta_{APP2}]/s$ ならば $t_0^* \rightarrow \infty$.

図6は定理5.2の内容を整理したものである。図5と比較すると、最適解の構造がかなり似通っていることが分かる。すなわち、 $E[\eta_{APP2}]$ と $r_{APP2}(0)$ をそれぞれ $E[\eta_{APP1}]$ と λ に置き換えることによって、定理5.2は定理5.1に帰着される。

次章では、本章で提案した近似手法とはまったく異なる概念を用いることによって、最適スリープ時間を近似的に導出することを試みる。

6. 拡散近似

待ち行列過程を近似するための有効な手法として、拡散近似がある^{8),9)}。拡散近似とは、待ち行列を構成する離散事象計数過程を連続状態拡散過程によって

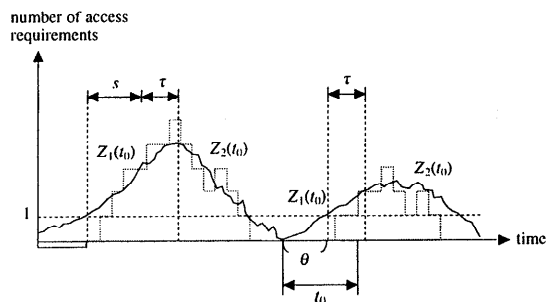


図7 拡散近似の概念図

Fig. 7 Schematic illustration of the diffusion approximation.

近似する手法である。この種の近似手法はコンピュータネットワークなどの複雑なシステムをモデル化するために利用されている。アクセス要求発生数の累積過程と $GI/G/1$ 待ち行列過程を、それぞれ拡散過程 $\{Z_i(t); t \geq 0\}$ ($i = 1, 2$) で近似することを考える(図7参照)。拡散過程 $Z_i(\cdot)$ ($i = 1, 2$) に対して、無限小平均 a_i (> 0) と拡散パラメータ b_i (> 0) を以下のように定義する。

$$a_i = \lim_{\Delta t \rightarrow 0} \frac{E[Z_i(t + \Delta t) - Z_i(t) | Z_i(t)]}{\Delta t}, \quad (47)$$

$$b_i = \lim_{\Delta t \rightarrow 0} \frac{\text{Var}[Z_i(t + \Delta t) - Z_i(t) | Z_i(t)]}{\Delta t}. \quad (48)$$

これより、確率過程 $\{Z_i(t); t \geq 0\}$ ($i = 1, 2$) はドリフト項を持つ1次元ブラウン運動過程となる。すなわち、 $t = 0$ からスリープ終了時刻を経てアクセス処理を開始するまでは $Z_1(t) = a_1 t + \sqrt{b_1} W(t)$ で(平均の意味で)増加し、アクセス処理開始からアクセス要求の数が0になるまで $Z_2(t) = Z_1(\theta) + a_2 t + \sqrt{b_2} W(t)$ で(平均の意味で)減少する。ここで、 $\{W(t), t \geq 0\}$ は平均0、分散 t の標準ブラウン運動過程である。図7における空き期間 θ が t_0 よりも小さい(スリープモードに入らない)場合も、 $s + \tau$ を τ に置き換えることにより同様に成り立つ。1次と2次モーメントのマッチングを考慮することによって、ただちに

$$\begin{aligned} a_1 &= \lambda (> 0), & b_1 &= \lambda^3 \sigma_a^2, \\ a_2 &= \lambda - \mu (< 0), & b_2 &= \lambda^3 \sigma_a^2 + \mu^3 \sigma_s^2 \end{aligned}$$

を得る。

拡散過程は連続状態確率過程であるため、システム内にアクセス要求が発生しないという事象を厳密に取り扱うことが困難となる。そこで、拡散過程 $Z_1(t)$ が初めて1に到達する時間、

$$\theta = \inf\{t \geq 0; Z_1(t) \geq 1 \mid Z_1(0) = 0\} \quad (49)$$

がサーバの空き期間に対応するものと考えれば、空き期間の確率分布関数 $I(x | t)$ を θ の確率分布関数によって近似することができる。明らかに、 θ の確率分布関数は逆ガウス分布¹⁰⁾となるので、空き期間分布 $I(x | t)$ は

$$\begin{aligned} I(x | t) &\approx I_{\text{dif}}(x) = \Pr\{\theta \leq x\} \\ &= \Phi\left[\frac{a_1 x - 1}{\sqrt{b_1 x}}\right] + \exp\left(\frac{2a_1}{b_1}\right) \\ &\quad \times \Phi\left[\frac{-a_1 x - 1}{\sqrt{b_1 x}}\right] \end{aligned} \quad (50)$$

となる。ここで、 $\Phi[\cdot]$ は標準正規分布関数である。これより、 η_t の期待値は

$$\begin{aligned} E[\eta_t] &\approx E[\theta] = 1/a_1, \\ E[\eta_t \wedge t_0] &\approx E[\theta \wedge t_0] \\ &= \frac{1}{a_1} \left\{ \Phi\left[\frac{-1 + a_1 t_0}{\sqrt{b_1 t_0}}\right] \right. \\ &\quad \left. - \exp\left(\frac{2a_1}{b_1}\right) \Phi\left[\frac{-1 - a_1 t_0}{\sqrt{b_1 t_0}}\right] \right\} \\ &\quad + t_0 \left\{ \Phi\left[\frac{1 - a_1 t_0}{\sqrt{b_1 t_0}}\right] \right. \\ &\quad \left. - \exp\left(\frac{2a_1}{b_1}\right) \Phi\left[\frac{-1 - a_1 t_0}{\sqrt{b_1 t_0}}\right] \right\} \end{aligned} \quad (51)$$

となる。上述の関係式を式 (6) と式 (13) に代入することによって、1 サイクルの期待時間の近似式は

$$\begin{aligned} T(t_0) &\approx T_{\text{dif}}(t_0) \\ &= \frac{1}{(1-\rho)\bar{I}_{\text{dif}}(t_0)} (\tau + E[\theta]) \\ &\quad + \frac{s}{1-\rho} \end{aligned} \quad (53)$$

となり、1 サイクル中に生じる総期待消費電力の近似式は

$$\begin{aligned} C(t_0) &\approx C_{\text{dif}}(t_0) \\ &= \frac{P_1}{\bar{I}_{\text{dif}}(t_0)} \left(\frac{\tau}{1-\rho} + \frac{\rho}{1-\rho} E[\theta] \right. \\ &\quad \left. + E[\theta \wedge t_0] \right) + \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} s \end{aligned} \quad (54)$$

となる。式 (52) と (53) を用いて、定常状態における単位時間あたりの期待消費電力の近似式は

$$V(t_0) \approx V_{\text{dif}}(t_0) = C_{\text{dif}}(t_0)/T_{\text{dif}}(t_0) \quad (55)$$

によって表現される。

いま、以下に示すような非線形関数

$$\begin{aligned} q_{\text{dif}}(t_0) &\equiv \left\{ (1-\rho)P_1 - r_{\text{dif}}(t_0)s[\rho P_1 \right. \\ &\quad \left. + (1-\rho)P_2] \right\} \bar{I}_{\text{dif}}(t_0) T_{\text{dif}}(t_0) \end{aligned}$$

$$+ r_{\text{dif}}(t_0)s\bar{I}_{\text{dif}}(t_0)C_{\text{dif}}(t_0) \quad (56)$$

を定義する。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、

$$r_{\text{dif}}(t) = \frac{dI_{\text{dif}}(t)/dt}{\bar{I}_{\text{dif}}(t)} \quad (57)$$

によって定義される。式 (49) から $r_{\text{dif}}(t)$ に関する以下の補題を得る。

補題 6.1 : 以下の諸量を定義する。

$$t_m \equiv -\frac{3b_1}{2a_1^2} + \frac{1}{a_1} \sqrt{1 + \frac{9b_1^2}{4a_1^2}}, \quad (58)$$

$$t_n \equiv \frac{2}{3b_1}. \quad (59)$$

このとき方程式

$$r_{\text{dif}}(t) = \frac{a_1^2}{2b_1} + \frac{3}{2t} - \frac{1}{2b_1 t^2} \quad (60)$$

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。

証明は文献 10) を参照されたい。以上の準備の下で、定常状態における単位時間あたりの期待消費電力の近似式 $V_{\text{dif}}(t_0)$ を最小にする最適スリープ時間 t_0^* に関する次の結果が得られる。

定理 6.2 : $\rho P_1 + (1-\rho)P_2 > V_{\text{dif}}(t_0)$ の下で $t_e \in (t_m, t_n)$ が存在すると仮定する。

(1) $q_{\text{dif}}(t_e) \leq 0$ かつ $q_{\text{dif}}(\infty) > 0$ のとき、(i) $q_{\text{dif}}(0) < 0$ ならば、定常状態における単位時間あたりの期待消費電力を最小にする有限で唯一の最適スリープ時間の近似値 $t_0^* \in (t_e, \infty)$ が存在し、非線形方程式 $q_{\text{dif}}(t_0^*) = 0$ を満足する。そのときの最小期待消費電力は、

$$V_{\text{dif}}(t_0^*) = \rho P_1 + (1-\rho)P_2 - \frac{(1-\rho)P_1}{r_{\text{dif}}(t_0^*)s} \quad (61)$$

となる。(ii) $q_{\text{dif}}(0) \geq 0$ ならば $q_{\text{dif}}(\bar{t}_0) = 0$ を満足する $\bar{t}_0 \in (t_e, \infty)$ が存在する。さらに、不等式

$$r_{\text{dif}}(\bar{t}_0) < \frac{P_1\{1 + \lambda s + \lambda \tau\}}{s\{P_2 - \lambda \tau(P_2 - P_1)\}} \quad (62)$$

が成立するならば、最適スリープ時間の近似値は $t_0^* = \bar{t}_0$ となり、そのときの最小期待消費電力は

$$V_{\text{dif}}(0) = \frac{P_1(\tau + \rho E[\theta]) + \{\rho P_1 + (1-\rho)P_2\}s}{\tau + E[\theta] + s} \quad (63)$$

によって与えられる。そうでない場合は $t_0^* = 0$ となる。

(2) $q_{\text{dif}}(t_e) > 0$ または $q_{\text{dif}}(\infty) \leq 0$ のとき、 $P_1 > \lambda s(P_2 - P_1)$ ならば $t_0^* = 0$ となり、そうでなければ $t_0^* \rightarrow \infty$ となる。そのときの最小期待消費電力は

$$V_{\text{dif}}(\infty) = P_1 \quad (64)$$

となる。

定理 6.2 による最適スリープ時間の性質は定理 5.1 と 5.2 によって得られたものよりもかなり複雑となることが分かる。しかしながら、拡散近似による解法はアクセス要求時間間隔の年齢特性 (DHR もしくは IHR) に依存しないという利点があることも事実であり、自動スリープシステムの時間的挙動を大域的にモデル化するためには非常に便利である。

次章では、アクセス要求時間間隔に指数分布を仮定できない場合に焦点を絞り、定常状態における単位時間あたりの期待消費電力を最小にする最適スリープ時間を数値的に求めるとともに、提案された 3 種類の近似手法を近似精度の観点から比較する。

7. 数 値 例

ここでは前報³⁾と同様に、アクセス要求時間を連続時間マルコフ連鎖においてある状態に吸収されるまでの時間と見なすことにより、以下のような位相型分布 (phase-type distribution) を仮定する。

$$F(t) = 1 - \alpha \exp(Tt)e. \tag{65}$$

ここで、 α は $m (> 0)$ の長さを持つ確率ベクトル、 e はすべての要素が 1 である長さ m の列ベクトル、 T は対角要素は負で非対角要素は正であるような $m \times m$ 行列であり、 Te の各要素は非正であるとする。また、アクセス処理時間分布にはパラメータ μ の指数分布を仮定する。

このとき、コンピュータのバッファ内に蓄積されるアクセス要求数はバケーションをとまなう PH/M/1 待ち行列を形成し、空き期間分布 $I(x|t)$ は

$$I(x|v) = \int_0^\infty \bar{F}(x+v+t)dG(t) + \int_0^\infty \int_0^{v+t} I(x|v+t-u) \times dF(u)dG(t) \tag{66}$$

となる。式 (65) の両辺を v に関してラプラス変換すると、

$$I^*(x|z) = \frac{I(x|0) - \mu\alpha(zI - T)^{-1} \exp(Tx)e}{z - \mu + \mu\alpha(zI - T)^{-1}e} \tag{67}$$

を得る。ここで、 z はラプラス演算子であり、 I は $m \times m$ の単位行列である。

また、 $I(x|0)$ はバケーションをとまなわない通常の GI/M/1 待ち行列における空き期間分布であり、 x に関するラプラス・スチルチェス変換 (Laplace-Stieltjes transform) は以下になることが知ら

れている¹¹⁾。

$$\int_{x=0}^\infty \exp(-ux)dI(x|0) = \frac{\mu\{f^*(u) - \delta\}}{\mu(1 - \delta) - u} \tag{68}$$

ここで、 $f^*(u)$ はアクセス要求発生時間間隔分布 $F(t)$ のラプラス・スチルチェス変換である。また、トラフィック強度が $\rho < 1$ のとき、 δ は以下の方程式を満足する 1 よりも小さい唯一の解となる。

$$\delta = f^*(\mu(1 - \delta)). \tag{69}$$

式 (67) の逆ラプラス変換をとることにより、空き期間分布 $I(x|v)$ を理論的に得ることができる。

数値計算において、以下のようなパラメータを設定する。

$$T = \lambda \begin{bmatrix} -2.22 & 0.44 \\ 0.22 & -0.44 \end{bmatrix}, \\ \alpha = (1.0, 0.0), \quad \lambda = 0.1, 0.2, \dots, 0.9, \\ \mu = 1.0.$$

また、他のモデルパラメータを $\tau = 0.05$, $P_1 = 1.0$, $P_2 = 3.0$, $s = 1.2$ とし、アクセス要求発生時間間隔のパラメータ λ を変化させた場合の最適スリープ時間と、その場合の最小期待消費電力を数値的に求める。近似精度の評価には、次のような相対誤差を用いる。

$$\frac{(\text{近似値}) - (\text{実際値})}{(\text{実際値})} \times 100 (\%).$$

図 8 は、スリープ時間 t_0 に対する定常状態における単位時間あたりの期待消費電力の振舞いを図示したものである。これより、近似 2 と拡散近似に基づいて算出された期待消費電力は、それぞれ実際の値よりもやや高めに見積もられていることが分かる。すなわち、アクセス要求時間間隔分布に位相型分布を仮定した場

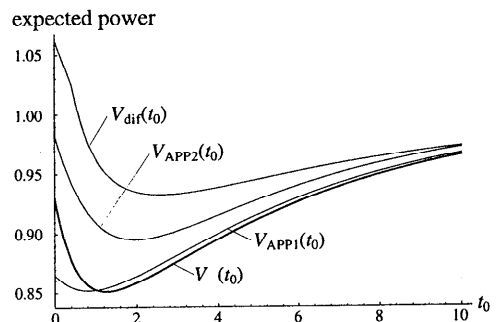


図 8 定常状態における単位時間あたりの期待消費電力の振舞い ($\lambda = 0.5, \mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$)

Fig. 8 Behaviour of expected power consumed per unit time in the steady-state ($\lambda = 0.5, \mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$).

表 1 近似手法の比較 ($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$)

Table 1 Comparison of approximation methods ($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$).

λ	exact		approximation 1			approximation 2			diffusion approximation		
	t_0^*	$V(t_0^*)$	t_0^*	$V_{APP1}(t_0^*)$	error (%)	t_0^*	$V_{APP2}(t_0^*)$	error (%)	t_0^*	$V_{dif}(t_0^*)$	error (%)
0.1	0.000	0.351	0.000	0.264	-24.690	0.000	0.375	6.750	0.000	0.392	11.741
0.2	0.000	0.558	0.000	0.475	-14.878	0.000	0.623	11.606	0.000	0.667	19.632
0.3	0.000	0.692	0.000	0.641	-7.408	1.645	0.760	9.775	2.008	0.812	17.308
0.4	0.000	0.787	0.441	0.766	-2.659	1.944	0.839	6.646	2.320	0.886	12.603
0.5	0.449	0.857	0.843	0.852	-0.602	1.988	0.895	4.443	2.599	0.932	8.799
0.6	0.644	0.910	1.071	0.912	0.162	1.980	0.936	2.812	2.880	0.962	5.654
0.7	0.793	0.950	1.234	0.953	0.295	1.979	0.965	1.573	3.182	0.980	3.154
0.8	0.955	0.978	1.384	0.979	0.194	2.000	0.984	0.709	3.517	0.991	1.374
0.9	1.185	0.994	1.582	0.994	0.065	2.112	0.996	0.199	3.898	0.997	0.337

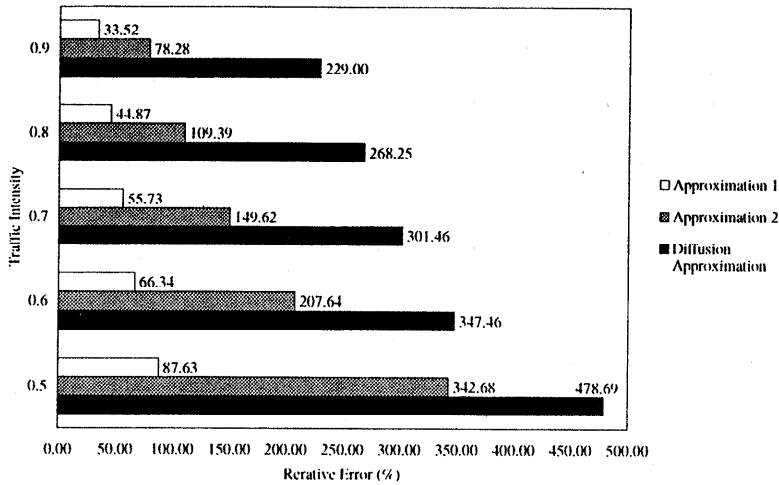


図 9 期待消費電力の相対誤差におけるトラフィック強度の従属性 ($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$)

Fig. 9 Dependence of traffic intensity in relative error of expected power consumed ($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$).

合, $V_{APP2}(t_0)$ と $V_{dif}(t_0)$ は関数 $V(t_0)$ の上界を形成しており, 若干悲観的な意思決定を行うことを示している。これに対して $V_{APP1}(t_0)$ は, スリープ時間が比較的大きい場合はかなりの精度で $V(t_0)$ を近似することができるが, 小さいスリープ時間に対しては楽観的な予測を行うことが分かる。表 1 は, 最小期待消費電力の近似精度を相対誤差の観点から比較したものである。パラメータ λ の値が大きくなれば, いずれの近似手法によって得られた結果もその精度を増すことが分かる。

一方, 図 9 はトラフィック強度の変化に対する各近似手法に基づいた相対誤差を比較したものである。この図から, トラフィック強度が小さくなるにつれて, 近似 2 と拡散近似に基づいて算出された最適スリープ時間の相対誤差は極端に増加する傾向にあり, 近似 1 が最も良好な近似結果を与えることが読み取れる。本論文で考察しているモデルでは, 一度コンピュータが

アクセスに対する処理を開始すると, バッファ内に蓄積しているアクセス要求がなくなるまで継続的にアクセス処理を行わなければならない。そのため, 実際にアクセスに対する処理を行う時間が, 前報³⁾で考察されたアクセス要求をキャンセルするモデルと比較して長くなっている。このようなモデルの構造上の相違点が, 近似手法の有効性に大きく影響を及ぼしていると考えられる。さらに, 図 9 において求められた誤差が相対的に大きい場合, 近似手法の有効性が懸念される場合も十分に考えられる。そこで, 近似的に導出された最適スリープ時間を期待消費電力 $V(t_0)$ に代入した結果を表 2 に示す。表の結果から, 最適スリープ時間の期待消費電力に対する誤差の影響は, 近似 1 で 0.1%程度であり, 十分実用に耐えうる近似であることが分かる。

次に, 実際の観測データに基づいた最適スリープ機能の設計事例について述べる。著者らの所属する広島

表2 定常状態における単位時間あたりの期待消費電力に関する相対誤差の評価
($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$)

Table 2 Evaluation of relative error on expected power consumed per unit time in the steady-state ($\mu = 1.0, \tau = 0.05, s = 1.2, P_1 = 1.0, P_2 = 3.0$).

λ	approximation 1			approximation 2			diffusion approximation		
	t_{APP1}^*	$V(t_{APP1}^*)$	error (%)	t_{APP2}^*	$V(t_{APP2}^*)$	error (%)	t_{dif}^*	$V(t_{dif}^*)$	error (%)
0.1	0.000	0.351	0.000	0.000	0.351	0.000	0.000	0.351	0.000
0.2	0.000	0.558	0.000	0.000	0.558	0.000	0.000	0.558	0.000
0.3	0.000	0.692	0.000	1.645	0.715	3.364	2.008	0.722	4.384
0.4	0.441	0.788	0.127	1.944	0.805	2.298	2.320	0.812	3.178
0.5	0.843	0.858	0.109	1.988	0.870	1.545	2.599	0.880	2.715
0.6	1.071	0.911	0.122	1.980	0.920	1.025	2.880	0.931	2.284
0.7	1.234	0.951	0.103	1.979	0.956	0.599	3.182	0.966	1.644
0.8	1.384	0.978	0.058	2.000	0.980	0.274	3.517	0.986	0.916
0.9	1.582	0.994	0.019	2.112	0.994	0.079	3.898	0.997	0.306

表3 定常状態における単位時間あたりの最小期待消費電力に関する実データを用いた評価 ($m_a = 0.618, \beta_a = 45.781, m_s = 0.149, \beta_s = 370.263, \tau = 0.0, P_1 = 1.0, P_2 = 10.0$)

Table 3 Evaluation of minimum expected power consumed per unit time in the steady-state using the real data ($m_a = 0.618, \beta_a = 45.781, m_s = 0.149, \beta_s = 370.263, \tau = 0.0, P_1 = 1.0, P_2 = 10.0$).

s	approximation 1		approximation 2		diffusion approximation	
	t_{APP1}^*	$V(t_{APP1}^*)$	t_{APP2}^*	$V(t_{APP2}^*)$	t_{dif}^*	$V(t_{dif}^*)$
1.0	0.000	0.846	0.000	0.850	0.000	0.058
2.0	0.000	0.859	0.000	0.867	0.000	0.882
3.0	0.000	0.871	0.000	0.883	0.000	0.905
4.0	0.000	0.884	0.000	0.899	0.000	0.928
5.0	0.000	0.896	0.860	0.914	29.421	0.945
6.0	0.000	0.908	4.952	0.929	35.669	0.955
7.0	0.000	0.919	12.062	0.942	41.571	0.963
8.0	2.341	0.931	22.168	0.953	47.164	0.970
9.0	8.040	0.942	35.342	0.962	52.475	0.976
10.0	16.907	0.951	51.707	0.971	57.530	0.982

大学工学部第二類計数工学教育科目に設置してある3台のワークステーションに対して、アクセス要求時間間隔と使用時間のデータを採集した。データの採集期間は3カ月であり、観測データ数はそれぞれ、アクセス要求時間間隔データ247個、使用時間データ226個である。文献3)における結果より、アクセス要求時間間隔分布とアクセス処理時間分布の年齢特性はともにDHR (decreasing failure rate) 特性を示すので、各々のデータに対する理論分布の適合度検定を行った。 χ^2 適合度検定によって標本が確率分布 $F(\cdot)$ および $G(\cdot)$ から採取されたという帰無仮説を上側5%有意水準で検定した結果、アクセス要求時間間隔にはワイブル分布、アクセス処理時間にはガンマ分布が適合していることが分かった。これより、アクセス要求時間間隔分布として形状パラメータ $m_a = 0.618$ と尺度パラメータ $\beta_a = 45.781$ のワイブル分布を、アクセス処理時間分布として形状パラメータ $m_s = 0.149$ と尺度パラメータ $\beta_s = 370.263$ のガンマ分布をそれぞれ

仮定する。

最適スリープ時間を数値的に求めることにより、実際のコンピュータシステムにおける省電力機能を定量的に評価する。使用されるモデルパラメータは以下のように設定した。

$$s = 1, 2, \dots, 10, \quad \tau = 0.0,$$

$$P_1 = 1.0, \quad P_2 = 10.0.$$

4章と5章において明らかにされたように、最適スリープ時間は電力比 P_2/P_1 に依存するため、ここでは簡単のため $P_1 = 1.0$ と設定した。表3は、パラメータ s の変化に対する最適スリープ時間の近似値および最小期待消費電力の近似値の従属性を調べたものである。たとえば、 $s = 10.0$ のとき、各近似に基づいた最小期待消費電力は約0.95~0.98ワットとなるが、スリープ機能を持たないコンピュータの単位時間あたりの期待消費電力 ($t_0^* \rightarrow \infty$ の場合) が1ワットになることを考慮すれば、通常に消費する総電力の約2~5%ほどの節電効果がある。一方、アクセス要求が0になった時

点でつねにスリープモードに移項した場合と比較すると、期待消費電力の近似値は $V_{APP1}(0) = 0.953$ ワット、 $V_{APP2}(0) = 0.987$ ワット、 $V_{dif}(0) = 1.052$ ワットとなり、それぞれ 0.2%、1.65%、7.13% の節電効果が期待できる。これらの結果から、本論文で考察したシステムは、バッファにおいてアクセス要求の待ちをとまわらないシステム³⁾よりも省電力効果が低くなるが、確実に節電の効果が実現されていることが分かる。

8. 結 論

本論文では、アクセス要求がバッファ内で待ち行列を形成するようなシステムに対して、定常状態における単位時間あたりの期待消費電力を定式化し、それを最小にするような最適スリープ時間を求める手続きを提案した。バッファ内におけるアクセス要求数の時間的推移が $M/G/1$ 待ち行列過程によってモデル化できる場合、最適スリープ時間は 0 か無限大となり、コンピュータのアイドル状態を検出するとただちにスリープモードに移行するか、もしくはまったくスリープモードに移行しないことが最適となる。また、アクセス要求数の時間的推移が $GI/G/1$ 待ち行列過程によってモデル化できる場合に対して、3 種類の近似解法を提案した。特に、アクセス要求時間間隔の均衡分布に基づいた近似手法（近似 1）は実利用上有効であり、本論文で得られた知見はコンピュータシステムの自動スリープ機能を省電力の観点から設計するうえで非常に有用であると結論づけられる。

今後の課題として、処理待ち時間等、他の性能評価尺度を考慮した自動スリープシステムの設計について考察する必要がある。また、マルチタスクの OS を搭載しているコンピュータシステムに接続されているハードディスクやネットワークプリンタ等からデータを収集することにより、実際のシステムに対する省電力効果をさらに詳細に検討する予定である。

謝辞 本研究の一部は文部省科学研究費奨励研究 (A) Grant No.09780411 および文部省科学研究費基盤研究 (C) Grant No.09898007 の助成の下で行われたものである。本論文に対して貴重な助言をいただいた 2 名の査読者にこの場を借りて感謝申し上げる。

参 考 文 献

- 1) 下野野 享：ノートブック PC のパワーマネジメント、情報処理、Vol.38, No.2, pp.135-142 (1997).
- 2) 相原 達、関家一雄、黒田昭裕：ノートブック PC におけるローパワーシステム技術、電子情報

通信学会誌、Vol.80, No.4, pp.370-376 (1997).

- 3) 岡村寛之、土肥 正、尾崎俊治：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 I—再生過程によるモデル化、情報処理学会論文誌、Vol.39, No.6, pp.1858-1869 (1998).
- 4) Sandoh, H., Hirakoshi, H. and Kawai, H.: An Optimal Time to Sleep for an Auto-Sleep System, *Computers & Oper. Res.*, Vol.23, No.3, pp.164-168 (1994).
- 5) Ross, S.M.: *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco (1970).
- 6) Takagi, H.: *Queueing Analysis*, Vol.1, Elsevier Sci. Pub., The Netherlands (1991).
- 7) Miyazawa, M.: A Generalized Pollaczek-Khintchine Formula for the $GI/GI/1/K$ Queue and Its Application to Approximation, *Commun. Stat. Stochastic Models*, Vol.3, pp.53-65 (1989).
- 8) Kobayashi, H.: Application of the diffusion approximation to queueing networks I, equilibrium queue distributions, *J. ACM.*, Vol.21, pp.316-328 (1974).
- 9) Kobayashi, H.: Application of the diffusion approximation to queueing networks II, nonequilibrium queue distributions and applications to computer modeling, *J. ACM.*, Vol.21, pp.459-469 (1974).
- 10) Chhikara, R.S. and Folks, J.L.: The inverse Gaussian distribution as lifetime model, *Technometrics*, Vol.19, pp.461-468 (1977).
- 11) Marshall, K.T.: Some relationships between the distributions of waiting time, idle time and interoutput time in the $GI/G/1$ queue, *SIAM J. Appl. Math.*, Vol.16, pp.324-327 (1968).

付 録

定理 4.1 の証明：以下のような記号を定義する。

$$AT_{M/G/1}(t_0) \equiv \exp(-\lambda t_0) T_{M/G/1}(t_0), \quad (70)$$

$$AC_{M/G/1}(t_0) \equiv \exp(-\lambda t_0) C_{M/G/1}(t_0). \quad (71)$$

式 (70) と (71) を用いて、定常状態における単位時間あたりの期待消費電力 $V_{M/G/1}(t_0)$ は

$$\begin{aligned} V_{M/G/1}(t_0) &= C_{M/G/1}(t_0)/T_{M/G/1}(t_0) \\ &= AC_{M/G/1}(t_0)/AT_{M/G/1}(t_0) \end{aligned} \quad (72)$$

によって表現される。さらに、 $V_{M/G/1}(t_0)$ の t_0 に関する 1 階導関数を求めると、

$$\begin{aligned} \frac{d}{dt_0} V_{M/G/1}(t_0) &= \frac{1}{AT_{M/G/1}(t_0)^2} \left\{ (P_1 + P_2 \lambda s \right. \\ &\quad \left. - \frac{\rho P_1 \lambda s}{1 - \rho}) \right\} \end{aligned}$$

$$\begin{aligned} & \times \exp(-\lambda t_0) AT_{M/G/1}(t_0) \\ & + AC_{M/G/1}(t_0) \\ & \times \left. \frac{\lambda s \exp(-\lambda t_0)}{1-\rho} \right\}. \quad (73) \end{aligned}$$

ここで、次のような非線形関数 $q_{M/G/1}(t_0)$ を定義する.

$$\begin{aligned} q_{M/G/1}(t_0) & \equiv \frac{AT_{M/G/1}(t_0)^2}{\exp(-\lambda t_0)} \frac{d}{dt_0} V_{M/G/1}(t_0) \\ & = \left(P_1 - \frac{\rho P_1 + (1-\rho)P_2}{1-\rho} \lambda s \right) \\ & \quad \times AT_{M/G/1}(t_0) + \frac{\lambda s}{1-\rho} \\ & \quad \times AC_{M/G/1}(t_0). \quad (74) \end{aligned}$$

この非線形関数 $q_{M/G/1}(t_0)$ の符号は $V_{M/G/1}(t_0)$ の1階導関数の符号と等しいので、関数 $q_{M/G/1}(t_0)$ を t_0 について微分すると、 $q_{M/G/1}(t_0)$ が定数となることが分かる. すなわち、

$$q_{M/G/1}(t_0) = \left\{ P_1 - \lambda s(P_2 - P_1) \right\} \frac{\tau + 1/\lambda}{1-\rho} \quad (75)$$

となる. ここで、 $q_{M/G/1}(t_0)$ の符号と $V_{M/G/1}(t_0)$ の1階導関数の符号が等しいので、 $P_1 - \lambda s(P_2 - P_1) > 0$ ならば $V_{M/G/1}(t_0)$ は単調増加関数となる. よって、 $V_{M/G/1}(t_0)$ を最小にする最適スリープ時間は $t_0^* = 0$ となる. 逆に、 $P_1 - \lambda s(P_2 - P_1) \leq 0$ ならば $V_{M/G/1}(t_0)$ は単調減少関数となり、最適スリープ時間は $t_0^* \rightarrow \infty$ となる. (証終)

(平成10年1月19日受付)

(平成10年11月9日採録)



岡村 寛之

昭和48年生. 平成7年広島大学工学部第二類卒業. 平成9年同大学大学院工学研究科システム工学専攻博士課程前期修了. 同年(株)CSK入社. 在学中、待ち行列理論、コン

ピュータシステムの性能評価の研究に従事.



土肥 正

昭和40年生. 平成元年広島大学工学部第二類卒業. 平成4年同大学大学院工学研究科システム工学専攻博士課程後期中途退学. 同年広島大学工学部助手. 平成8年同助教授.

博士(工学). 平成4年ブリティッシュ・コロンビア大学(カナダ)客員研究員. 主として、信頼性理論および数理システムの研究に従事. 日本OR学会, 日本応用数理学会, 電子情報通信学会, システム制御情報学会, 計測自動制御学会, Society of Economic Dynamics and Control 各会員. *Communications on Dependability and Quality Management in Engineering of International Editorial Member.*



尾崎 俊治 (正会員)

昭和16年生. 昭和45年京都大学大学院工学研究科数理工学専攻博士課程修了. 同年広島大学工学部経営工学科講師, 同年同助教授, 昭和61年広島大学工学部第二類教授. 工

学博士. 主に、応用確率論, 信頼性工学およびシステム工学の研究に従事. 昭和45~47年南カリフォルニア大学(アメリカ)客員研究員, 昭和51~52年マンチェスター大学(イギリス)客員研究員. 日本OR学会第4回大西文献賞受賞. 著書「品質管理・信頼性のための統計分布ハンドブック」, 「確率モデル入門」, 「Stochastic System Reliability Modeling」, 「Applied Stochastic System Modeling」等. 編著「Stochastic Models in Reliability Theory」, 「Reliability Theory and Applications」等. 日本OR学会フェロー, 日本経営工学会, 電子情報通信学会, システム制御情報学会各会員. *RAIRO - Operations Research, Computers & Mathematics with Applications, Applied Stochastic Models and Data Analysis, Reliability, Quality and Safety Engineering, IIE Transactions on Quality and Reliability Engineering*等の Editorial Member.