タイム・シェア処理における多重度*の効果†

山崎源治†† 逆瀬川浩孝†††

ジョブのシステムへの到着過程が Poisson 過程である Limited Processor Sharing System (LiPS システム) を考える．この LiPS システムはタイム・シェア処理システムの一般化であり，その特別な場合として，待ち行列ネットワークでの Processor Sharing 規律，通常の有限多重度のタイム・シェア処理方式などを含んでいる．本稿では，この LiPS システムの平均系内ジョブ数の近似式・不等式を導く．この不等式は，ジョブのサービス要求量の変動係数 (C_s) が1より小さいときは，多重度を増すことはシステム効率を減少させること，$C_s>1$ のときはその逆となること，を示唆している．さらに，得られた近似解と厳密解の比較から，その精度が非常によいことが明らかになる．上述の不等式・近似式は，共に点過程の理論から導かれる．すなわち，最初に点過程の保存則を基に，一種の'平衡式'を導き，それから平均系内ジョブ数の厳密解を求める．この解は，ある未知量を含んでいるが，この未知量に対して適当な評価及び適当な仮定を設けることにより，不等式・近似式が導かれる．それゆえ，得られた近似式はその簡便性と合わせて，どこまで厳密解で，どこをどのように近似しているかが明確である，という特長を持っている．

1. はじめに

計算機システムの性能評価モデルとして，網型待ち行列モデルが用いられることが多い．このモデルにおいては，各資源・各ジョブはランダムな要因を含んでいてもよく，それでいて，いわゆる積形式解という陽解を持つので，実際に近い解析ができるという利点がある．しかし，一方で，解析を容易にするために，いくつかの制約が必要となっている．時分割処理に関するものもその制約の一つである．ラウンドロビン方式と呼ばれる時分割処理方法は，一定数のジョブに対してタイムスライスを割り当て，それらのジョブを決められた順番で一つずつ処理して行く方法である．性能評価モデルにこの処理方法を忠実にモデル化したものを取り入れると解析が難しくなるため，網型待ち行列モデルでは，資源分割 (Processor Sharing) 規律によって近似している．PS 規律とは，オーバヘッド時間を無視し，多重度制限を考えにいれないラウンドロビン方式で，そのタイムスライスを限りなく短くしたときの極限状態として定義されている．これはまた，ジョブ数に応じて機能が低下する複数サーバと考えてもよい．本稿では，この PS 規律を現実に近付けるために，多重度制約を設けた PS 規律を考え，ジョブがポアソン到着する単一サーバモデルで，この新しい規律の特性を調べる．

多重処理を行う場合に，多重度をいくつにすれば効率が最大になるのか，待ち時間が最小になるのか，という問題は非常に興味のあるところであるが，理論的にはほとんど解析されていない．ここで提案される多重度に制約のある (Limited) PS 規律（これを LiPS 規律と呼ぶ）は，上述したように通常の複数サーバシステムとよく似ている．多重度が m の LiPS 規律を持つシステム（これを m-LiPS システムと呼ぶ）と，mサーバシステムの違いは，サービス能力の配分の仕方にある．全稼働しているときのシステムのジョブ処理能力を1としたとき，前者ではシステム内のジョブ数の如何にかかわらず，常に1の能力で処理される．一方後者では，m人のサーバに対して能力を固定的に等分に割り振るため，システム内のジョブ数がmより小さい場合は，ジョブを割り当てられていないサーバの能力は使われないことになりシステムとしての処理能力は1より小さくなる．システム内のジョブ数がmより大きい場合は両者の振る舞いは同じものになる．それゆえ，最適多重度問題を考えるうえで，通常の複数サーバシステムにおける最適化問題が参考になると思われる．

待ち行列理論におけるサーバ数に関する最適化問題は，Stidham[4]，Brumelle[1] などにより議論されてきている．それらによれば，もし客のサービス時間がアーラン分布に従うならば，単一サーバシステムでのシステム内客数の分布は，利用率を等しくした複数サーバシステムでのそれよりも確率的に小さくなること[4]，また，もし客のサービス時間の変動係数が1よ

† The Effect of Multiplicity in Limited Processor Sharing Systems by GENJI YAMAZAKI (Tokyo Metropolitan Institute of Technology) and HIROTAKA SAKASEGAWA (Institute of Socio-Economic Planning, University of Tsukuba).

†† 東京都立科学技術大学管理工学科
††† 筑波大学社会工学系

* この多重度は，システムでのジョブの'最大同時処理数'を意味する．実際の計算システムでは，同時に処理しているジョブ数に応じて，システムの処理能力は変わるが，本稿ではそれは不変とする（詳細は第2章）．

り小さいならば，定常状態における単一サーバシステムでのシステム内客数の平均は，複数サーバシステムでのそれより小さくなること[1]，等が明らかになった．一方，変動係数が1より大きい場合には，複数サーバシステムでのシステム内ジョブ数の平均が単一サーバシステムでのそれよりも小さい例があり[1]，最適化問題はいまだに解決を見ていない．

著者らはすでに文献 6)において，LiPS システムの多重度に関する最適化問題を扱い，以下の理論結果を得ている．すなわち，もしジョブのサービス要求量の分布が NWU (変動係数が1より大きい，定義は次章でなされる) ならば LiPS システムのほうがシステム効率が高く，アーラン分布 (変動係数が1より小さい) ならば LiPS システムのほうがシステム効率が低くなる．ここで，システム効率の尺度はシステム内ジョブ数である．このように，LiPS システムにおいても上述のサーバ数に関する最適化問題同様，ジョブのサービス要求量の変動係数が重要な役割を果たすであろうことが予想できる．

本稿の第一の目的はジョブがポアソン到着する LiPS システムに限定すると，かなり広いサービス要求量の分布のクラス (NWUE・NBUE, 定義は次章でなされる) に対して

> 「変動係数が1より大きい (小さい) ならば，多重処理したほうがシステム効率は高く (低く) なる」

という事実が成立することを証明することである．この証明は，点過程論における保存則を用いて得られるシステム内ジョブ数の厳密解に基づいている．実用上は，与えられた多重度を持つシステムの平均システム内ジョブ数を具体的に求めることが必要となるが，上で得られる厳密解は計算が容易でない項を含んでいる．本稿の第二の目的は，この計算困難な厳密解に対して近似式を与えることである．多くの場合，待ち行列理論における近似式は数値実験に基づいた実験式で与えられるが，ここでは，指数分布を仮定したときに成り立ち，他の分布に対しても十分に妥当性があると思われる確率論的な仮定を使って厳密解の一部を置き換えることによって近似式を求めている．このような近似式の求め方は，実験式に基づくものに比べて，近似の意味づけが明確であるという特徴を持っている．得られた近似式は非常に簡単な形ではあるが，その精度は極めてよいということが，一連の厳密解との比較で明らかにされる．

本稿の構成は次のようになっている．次章では LiPS システムの定義，使われる記号，仮定などが与えられる．また，文献 6)で得られている主要な結果を述べる．3章では，点過程論における保存則を用いてシステム内ジョブ数の平均の厳密解を求め，それに基づいて不等式を導く．4章では，システム内ジョブ数の平均の近似式が与えられる．5章において，平均システム内ジョブ数が多重度およびサービス要求量の分布の変動係数によってどのような影響を受けるかを数値的に調べる．また，提案された近似式の精度が解析結果および数値計算の結果と比べられる．

2. モデルの記述，および主要な結果

本稿では，次のような待ち行列システムを扱う．ジョブはパラメータ λ のポアソン過程に従って到着し，それぞれが分布 $B(x)$ に従うサービス要求量を持ってくるものとする．一般にサービス要求量を S と書き，$B(x)$ の平均，2次モーメント，変動係数を $E[S]$, $E[S^2]$, C_s と書く．また，$\rho=\lambda E[S]$ と置く．システムは1から順に番号付けられた無数のセルからなり，各セルにはジョブが一つだけ入ることができるものとする．到着したジョブはいずれかのセルに入り，必要ならばセル間を移動しつつシステムに滞在し，サービス終了と同時にセルから退去するものとする．n 番目のセルを「ポジション n」と呼ぶ．システム内に n ジョブ存在するとき，到着したジョブはポジション $n+1$ にはいる．また，ポジション j のジョブが退去したときは，ポジション $j+1$, $j+2, \cdots$ のジョブは順にポジション $j, j+1, \cdots$ へ移る．

m をある自然数とし，この m に対して $\{r_j(n\,;m)$, $j=1,2,\cdots,n\,;n=1,2,\cdots\}$ を次のように定義する．

[1] $r_j(n\,;m)\geqq 0$, for all j,n.

[2] $\sum_j r_j(n\,;m)=1$, for all n.

[3] $r_j(n\,;m)=0$, for $j>m$.

この $\{r_j(n\,;m)\}$ を用いて，次のサービス規律を設ける．

> 「システムのジョブ処理能力は常に一定である．システム内のジョブ数が n の時，処理能力の $100\times r_j(n\,;m)$ パーセントをポジション j に振り向ける」

というものである．上述の待ち行列システムがこの規律に従ってジョブのサービスを行うとき，ある固定した m に対して同時にサービスを受けることができるのは，たかだかポジション1からポジション m にいる

ジョブだけである．また，処理能力一定のサーバが複数のジョブを同時にサービスする仕方を待ち行列論では PS 規律と呼んでいるので，この規律を m-Limited PS (m-LiPS) 規律と呼び，この規律に従ってジョブを処理する待ち行列システムを，m-LiPS システムと呼ぶ．

m-LiPS システムで，ジョブの到着過程，サービス要求量の分布を明示したい時は，通常の記号にならい，A/B/1 (m-LiPS) で表す．その際，一般分布を，G，または GI を用いて表すが，前者の場合はジョブ間の独立性を仮定せず，後者の場合のみ，それを仮定する．G/G/1 (1-LiPS) は FCFS 規律に従う通常の単一サーバシステムである．G/G/m で手空きのサーバは他のサーバを手伝いに行くというサービス規律もこの m-LiPS 規律によって表現できる．また PS 規律は $r_j(n; m)=1/n$ で m を無限大にした場合である．定義では m は自然数としていたが，便宜上，以下では $m=\infty$ も許すことにする．タイムスライスが一定のラウンドロビン方式は $r_j(n; m)=1/n$ $(n\leqq m)$ によって近似的にモデル化できる．この m-LiPS システムについて，次の記号を用いる．

$J(t; m)$：時刻 t におけるシステム内ジョブ数，

$V(t; m)$：時刻 t における各ジョブの残りサービス要求量の総和．

また，定常状態において，任意時点にシステムに n 個のジョブが存在する確率を p_n と書く．

二つの確率変数 X, Y に対して，次の記号を用いる．

$X \overset{d}{=} Y$；X と Y は確率的に等しい，すなわち，

$P\{X\leqq x\}=P\{Y\leqq x\}$, for all x.

$X \overset{d}{\leqq} Y$；X は Y より確率的に小さい，すなわち，

$P\{X>x\}\leqq P\{Y>x\}$, for all x.

非負の確率変数 X の分布関数を $F(x)$ とし，$1-F(x)$ を $\bar{F}(x)$ と書く．この $F(x)$ と任意の数 $\tau\geqq 0$ に対して次の分布関数 $F_\tau(x)$ を定義し，この $F_\tau(x)$ に従う確率変数を X_τ と書く．

$$1-F_\tau(x)=\bar{F}(x+\tau)/\bar{F}(\tau)$$

分布関数のクラスとして以下のものを定義する．

定義 $F(x)$ が NBU (New Better than Used) であるとは $X_\tau \overset{d}{\leqq} X$ が成り立つことである．$F(x)$ が NBUE (New Better than Used in Expectation) であるとは $E[X_\tau]\leqq E[X]$ が成り立つことである．

この定義で，不等号の向きが反対の時，それぞれ，$F(x)$ は NWU，NWUE であるという．NBU ならば NBUE であり，NWU ならば NWUE となる．NBUE，NWUE と変動係数の関係については，次の補題が成り立つ（例えば Stoyan[5] 参照）．

補題 1 $F(x)$ が NBUE (NWUE) ならばその変動係数は 1 より小さい（大きい）．

G/G/1 (m-LiPS) システムに関して著者らによって以下の事実が証明されている[6]．

補題 2 任意の m に対して，$V(t; m) \overset{d}{=} V(t; 1)$ $(t\geqq 0)$.

補題 3 m-LiPS システムにおける平衡条件は 1-LiPS システムにおけるそれと一致する．

この補題によれば，m-LiPS システムの平衡条件は $\rho<1$ である．以下，本稿では，定常状態のみを考える．

3. M/GI/1(m-LiPS) における不等式

本章ではまず，点過程の基礎理論に基づき，システム内ジョブ数に関する一種の平衡式を立て，それを使って $E[J(0; m)]$（これを $L(m)$ と表す）の厳密解を導く．次に，サービス要求量の分布をあるクラスに限定することにより，そのクラスの性質を使って厳密解から $L(m)$ と $L(1)$ に関する不等式を導く．

$\{N_0(t)\}$ をジョブの到着過程，$\{N_1(t)\}$ をジョブの退去過程，$\{X(t)\}$ を $\{N_0(t), N_1(t)\}$ の点以外で連続で，各点で右微分可能な標本関数を持つ確率過程（例えば，Doob[2]参照）とし，これらは同時に定常であるものとする．

$$\lambda=E[N_0((0,1])]=E[N_1((0,1])]$$

と書く．時刻 0 にあるジョブが到着（退去）したという条件の下で確率を考えるとき，$P_0(P_1)$ という記号を使い，$P_0(P_1)$ に関する期待値を $E_0(E_1)$ で表す．この時，次の補題が成り立つ (Miyazawa[3]の Corollary 3.1)．

補題 4 $E|X^+(0)|$, $E_0|X(0-)-X(0+)|$, $E_1|X(0-)-X(0+)|$ がいずれも有限ならば

$$E[X^+(0)] =\lambda E_0[X(0-)-X(0+)]+\lambda E_1[X(0-)-X(0+)] \tag{3.1}$$

ここで，$X^+(0)$ は右微分を表す．

さて，$X_n(t)$ として次のような確率過程を考える．

$$X_n(t)=I_{\{J(t;\, m)=n\}} \exp\{-\theta V(t; m)\} \tag{3.2}$$

ただし，ここで I_A は A が真であるとき 1，さもなければ 0 という値を取る確率変数であ

るとする．

この $X_n(t)$ は上記補題の仮定を満足しているので，これに補題4の結果を適用すると次の定理が得られる．

定理1 M/GI/1 (m-LiPS) では次の式が成立する．

$$\theta\varphi_1(\theta)p_1=\lambda\{1-\tilde{B}(\theta)\}p_0+\lambda\{\varphi_1(\theta)-\phi_1(\theta)\}p_1$$
$$\theta\varphi_n(\theta)p_n=\lambda\{\phi_{n-1}-\tilde{B}(\theta)\varphi_{n-1}(\theta)\}p_{n-1}+\lambda\{\varphi_n(\theta)-\phi_n(\theta)\}p_n,\ n=2,3,\cdots \quad (3.3)$$

ただし，$\tilde{B}(\theta)$ は $B(x)$ のラプラス変換

$\varphi_n(\theta),\phi_n(\theta)$ は $V(0;m)$ の条件付きラプラス変換で

$$\varphi_n(\theta)=E[e^{-\theta V(0;\,m)}\,|\,J(0;m)=n]$$
$$\phi_n(\theta)=E_1[e^{-\theta V(0;\,m)}\,|\,J(0+;m)=n]$$

によって定義される．

証明．m-LiPS システムでは $V(t;m)$ の標本関数は0でなければ，不連続点（到着時点）を除き傾きが -1 の直線になる．また，ジョブがポアソン到着するシステムを考えているので，E と E_0 は同じものを表す．さらに，定常状態におけるシステム内ジョブ数の分布は，任意時点でも退去直後時点でも同じになる．このことから，補題4の式の各項は以下のようにして求めることができる．

$$E[X_n(0)]=E[E[X_n^+(0)\,|\,J(0;m)]]$$
$$=p_n\frac{\mathrm{d}^+}{\mathrm{d}\theta}E[X_n(0)\,|\,J(0;m)=n]$$
$$\left(*\ \frac{\mathrm{d}^+}{\mathrm{d}\theta}\text{は右微分}\right)$$
$$=\theta p_nE[e^{-\theta V(0;\,m)}\,|\,J(0;m)=n]$$
$$E_0[X_n(0-)]=E_0[E_0[X_n(0-)\,|\,J(0-;m)]]$$
$$=p_nE[e^{-\theta V(0;\,m)}\,|\,J(0;m)=n]$$
$$E_0[X_n(0+)]=E_0[E_0[X_n(0+)\,|\,J(0+;m)]]$$
$$=p_{n-1}E[e^{-\theta(V(0;\,m)+S)}\,|\,J(0;m)=n-1]$$
$$=p_{n-1}\tilde{B}(\theta)E[e^{-\theta V(0;\,m)}\,|\,J(0;m)=n-1]$$
$$E_1[X_n(0-)]$$
$$=p_{n-1}E_1[e^{-\theta V(0;\,m)}\,|\,J(0+;m)=n-1]$$
$$E_1[X_n(0+)]=p_nE_1[e^{-\theta V(0;\,m)}\,|\,J(0+;m)=n] \quad (3.4)$$

これらの式を式(3.1)に代入すると式(3.3)が得られる．

一般に GI 型待ち行列システムを解析する際には，補助変数を用いてマルコフ過程を作り出し，その定常方程式を解析することが多いが，上の式はその定常方程式をラプラス変換したものにほかならない．補助変数法は方程式をたてる際かなり複雑であるのにたいして，本章の方法は単純であり，$X(t)$ の定義の仕方によっても多様な使い方が可能である．

定理1を使ってシステム内ジョブ数の平均を計算することができる．

定理2 M/GI/1(m-LiPS) におけるシステム内ジョブ数の平均 $L(m)$ は以下の式で与えられる．

$$L(m)=\frac{\rho}{1-\rho}+\frac{1}{1-\rho}\{L_q(1)-L^*\}$$
$$=L(1)+\frac{\rho}{1-\rho}\left\{L(1)-\frac{1}{\rho}L^*\right\} \quad (3.5)$$

ただし，$L^*=\lambda\sum_{n=1}^{\infty}p_nE_1[V(0;m)\,|\,J(0+;m)=n]$

また，$L_q(1)$ は M/GI/1 (1-LiPS) での平均待ちジョブ数を表し，ポラツェック・ヒンチンの公式を使って

$$L_q(1)=\frac{1+C_s^2}{2}\frac{\rho^2}{1-\rho}$$

と表される．

証明．式(3.3)の両辺に z^n をかけてすべての n について和を取ると

$$\theta\Phi(\theta,z)=\lambda z\{1-\tilde{B}(\theta)\}p_0+\lambda\Phi(\theta,z)-\lambda\Psi(\theta,z)+\lambda z\Psi(\theta,z)-\lambda z\tilde{B}(\theta)\Phi(\theta,z) \quad (3.6)$$

ただし，$\Phi(\theta,z)=\sum_{n=1}^{\infty}\varphi_n(\theta)p_nz^n$

$$\Psi(\theta,z)=\sum_{n=1}^{\infty}\phi_n(\theta)p_nz^n$$

ここで θ として $\lambda(1-z)$ とすれば，

$$\Phi(\theta,z)=\frac{\theta\Psi(\theta,z)}{\lambda z\{1-\tilde{B}(\theta)\}}-p_0$$

この両辺を z について微分して z を1に近付ければ (すなわち，$\theta\to0$)，

$$\text{左辺}=\lambda\sum E[V(0;m)\,|\,J(0;m)=n]p_n+L(m)$$
$$=L_q(1)+L(m)\quad(\because\ \text{補題2}) \quad (3.7)$$

右辺は L'Hospital の定理を使う．

$$\text{右辺}=\frac{1}{\rho}\left\{L(m)+\lambda\sum p_nE_1[V(0;m)\,|\,J(0+;m)=n]-\rho-\frac{1+C_s^2}{2}\rho^2\right\} \quad (3.8)$$

これらの式から式(3.5)を得る．

式(3.5)の右辺第二項は一般には計算がむずかしいが，サービス分布を制限すると上下限を評価することができる．

定理3 M/GI/1(m-LiPS) でジョブのサービス要求量の分布が NBUE (NWUE) ならば

$$L(m)\underset{(\leq)}{\geq}L(1). \quad (3.9)$$

証明　あるジョブの退去直後に，システム内に残されたポジション j にいるジョブがそれまでに受けたサービス量を X_j, 残余サービス量を Y_j とする．もし $B(x)$ がNBUE ならば，

$$E[Y_j]\leqq E[S] \tag{3.10}$$

であることから

$$\begin{aligned} &E_1[V(0;m)|J(0+;m)=n]\\ &\quad=E_1[E_1[\sum Y_j|J(0+;m)=n,\underline{X}]]\\ &\quad\leqq nE[S],\ \underline{X}=(X_1,X_2,\cdots,X_n), \end{aligned} \tag{3.11}$$

が成り立つ．これを式(3.5)に代入すれば，式(3.9)を得る．もし，$B(x)$ が NWUE ならば，不等式(3.10)の不等号は逆向きになるので，不等式(3.11)および(3.9)の不等号を逆向きにしたものが得られる．

この定理と補題1により，ジョブがポアソン到着するシステムにおいて多重処理がシステム効率を高くするか否かは，かなり一般的なサービス要求量の分布に対して，その変動係数だけで決まってしまうことが分かる．

4. M/GI/1/(m-LiPS) の平均システム内ジョブ数の近似式

前章では，$r_j(n;m)$ について制約のない M/GI/1/(m-LiPS) における平均システム内ジョブ数の定性的性質を調べた．その結論はそれ自身として興味のあるものであったが，実用上は，具体的な $r_j(n;m)$ が与えられたときの $L(m)$ の値を知ることが重要となろう．しかし残念ながら，$L(m)$ を正確に算定することは M/GI/m システムを解くのと同じ難しさがあるため，現時点では一般的に解くことはできない．そこで，本章では，$L(m)$ の近似式を考える．$L(m)$ は$r_j(n;m)$ に依存すると考えられるが，ここでは $r_j(n;m)$ として均等割当を仮定する．すなわち

$$\begin{aligned} r_j(n;m)&=1/n,\ \ j\leqq n\leqq m,\\ &\quad 1/m,\ \ j\leqq m\leqq n. \end{aligned} \tag{4.1}$$

とする．

式(3.5)の右辺第二項は退去時点における各ジョブの残余サービス要求量を含んでおり，一般の分布に対してこれを計算することは容易でない．そこで到着，退去時点でサービス中のジョブの平均残り仕事量は，それらのジョブの数にはよらず，再生理論で用いられる平均残存寿命で置き換えることにする．このように考えると，次の二つの近似的な関係式が得られる．以下これらの式を基に式(3.5)を変形して行く．

$$\begin{aligned} &E_1[V(0;m)|J(0+;m)=n]\\ &\quad\fallingdotseq\min(n,m-1)\frac{E[S^2]}{2E[S]}+\max(0,n-m+1)E[S] \end{aligned} \tag{4.2}$$

$$\begin{aligned} &E[V(0;m)|J(0;m)=n]\\ &\quad\fallingdotseq\min(n,m)\frac{E[S^2]}{2E[S]}+\max(0,n-m)E[S] \end{aligned} \tag{4.3}$$

式(4.2)はあるジョブが退去した時，もし待っているジョブがあればサービスが新たに開始され，残りのたかだか $m-1$ 個のジョブは残余サービス（その平均は $E[S^2]/2E[S]$）のサービスを継続する，ということを表している．また式(4.3)は任意時点でシステム内にジョブが n 個あればたかだか m 個のジョブがサービス中であることを表している．(4.1)のもとでは，式(4.2)，(4.3)は $m=\infty$ のとき等号で成り立つ（この直接的な証明は山崎[7]）．これが(4.1)の場合を扱う理由の一つである．

式(3.5)の第一式の右辺は式(4.2)，(4.3)を用いて以下のように変形できる．

$$\begin{aligned} L_q(1)-L^*&\fallingdotseq\lambda\sum p_n\{E[V(0;m)|J(0;m)=n]\\ &\qquad-E_1[V(0;m)|J(0+;m)=n]\\ &=\lambda\sum_{n=m}^{\infty}p_n\left(\frac{E[S^2]}{2E[S]}-E[S]\right)\\ &=\rho\left(\frac{C_s^2}{2}-1\right)\sum_{n=m}^{\infty}p_n \end{aligned} \tag{4.4}$$

一方，(3.3)の $n=1,2,\cdots,m$ について，両辺を θ で割り $\theta\to0$ として L'Hospital の定理，および，式(4.2)，(4.3)を用いることにより，次式を得ることができる．

$$p_n=\rho p_{n-1}\ \ (n=1,2,\cdots,m-1) \tag{4.5}$$

この式と，$p_0=1-\rho$, および確率条件とから式(4.4)の右辺の未知量が次のように求まる．

$$\sum_{n=m}^{\infty}p_n=\rho^m \tag{4.6}$$

以上の展開により，$L(m)$ の近似式として次の式が得られる．

$$\begin{aligned} L(m)&\fallingdotseq\frac{\rho}{1-\rho}\left\{1+\frac{C_s^2-1}{2}\rho^m\right\}\\ &=L(1)\left\{1+\frac{1-C_s^2}{1+C_s^2}(1-\rho^m)\right\} \end{aligned} \tag{4.7}$$

右辺の量を $L_a(m)$ と表す．$L_a(1)$ が $L(1)$ に等しいことはポラツェク・ヒンチンの公式を使えば容易に確かめることができる．また，サービス要求量の分布がフェーズ2のアーラン分布の場合には，確率母関数の方法を使って $L_a(2)$ と $L(2)$ が等しいことを確かめる

ことができる．さらに，$L_a(\infty)$ も式(4.2)，(4.3)が等号で成り立つことから，$L(\infty)$ に等しい．このように $L_a(m)$ は，近似式と言っても m のいくつかの特別な値については真値を与えている．このことと，サービス要求量の分布が極端に偏ったものでない限り，近似式(4.2)，(4.3)を仮定することは妥当なものと考えられるので，中間の m についてもかなり良い精度を持つことが期待できる．次章において，この近似式の精度について，数値的に考察する．

5. 数　値　例

本章では，$L(m)$ の厳密解が計算できる m-LiPS システムのいくつかの例を取り上げ，多重処理の効果を定量的に考察する．また，前章で提案された近似式 $L_a(m)$ の精度を調べる．

サービス要求量の分布の変動係数が1より大きい場合の例として，超指数分布を取り上げる．その密度関数は次のように表され，

$$f(t)=2r^2\mu e^{-2r\mu t}+2(1-r)^2\mu e^{-2(1-r)\mu t}\left(\frac{1}{2}\leqq r\leqq 1\right) \tag{5.1}$$

変動係数は r の関数として次式のように与えられる．

$$C_s=1+\frac{(1-2r)^2}{2r(1-r)} \tag{5.2}$$

$L(2)$ については，確率母関数を計算することにより陽に求めることができて，

$$L(2)=L(1)-(K-1)\rho R \tag{5.3}$$

$$\text{ただし，}K=\frac{1+C_s^2}{2},$$

$$R=1-\frac{2(1-\rho)}{2-\rho+K\sqrt{(1-\rho)^2+2\rho/K}-K(1-\rho)}$$

となる．しかし $m>2$ については解析的に解くことはむずかしいので，平衡方程式をたて，計算機によって $L(m)$ を計算することが必要である．変動係数と多重度が平均システム内ジョブ数にどのような影響を与えるかを調べるために，計算結果のいくつかを数表として与える(**表1**)．この表によれば，変動係数が大きくなるほど多重処理の効果が顕著であることが分かる．例えば，ρ が比較的小さい場合，多重処理をしない場合 ($m=1$) と，多重度2でも多重処理をした場合とでは，平均システム内ジョブ数が倍近くも違う例がある．その理由をきちんと説明することは難しいが，直観的には次のように考えれば理解できる．変動係数が大きくなると，極端に大きなサービス要求量を持つジョブが到着する可能性が大きくなるため，もし多重処理をしていないと，そのようなジョブのサービス中にジョブの待ち行列が長くなる可能性がある．もし多重処理をしていれば，そのジョブをサービスしながら同時に後から来るジョブに対してサービスできるので，大きなサービス要求量を持つジョブがボトルネッ

表 1 $M/H_2/1(m\text{-LiPS})$ における平均システム内ジョブ数
Table 1 Mean queue length $L(m)$ in $M/H_2/1(m\text{-LiPS})$.

(1) $C_s^2=2$

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.117	0.493	1.250	3.150	13.050
2	0.112	0.445	1.114	2.889	12.635
3	0.111	0.433	1.053	2.713	12.263
4	0.111	0.430	1.025	2.593	11.929
5	0.111	0.429	1.012	2.512	11.630
6	0.111	0.429	1.006	2.456	11.361
7	0.111	0.429	1.003	2.418	11.120
8	0.111	0.429	1.001	2.391	10.904
9	0.111	0.429	1.001	2.373	10.710
10	0.111	0.429	1.000	2.361	10.536
∞	0.111	0.429	1.000	2.333	9.000

(2) $C_s^2=4$

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.128	0.621	1.750	4.783	21.150
2	0.112	0.473	1.331	3.979	19.890
3	0.111	0.440	1.150	3.445	18.763
4	0.111	0.432	1.069	3.086	17.753
5	0.111	0.429	1.032	2.844	16.849
6	0.111	0.429	1.015	2.681	16.038
7	0.111	0.429	1.007	2.570	15.312
8	0.111	0.429	1.003	2.495	14.661
9	0.111	0.429	1.002	2.444	14.077
10	0.111	0.429	1.001	2.409	13.554
∞	0.111	0.429	1.000	2.333	9.000

(3) $C_s^2=6$

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.139	0.750	2.250	6.417	29.250
2	0.113	0.501	1.541	5.056	27.137
3	0.111	0.446	1.241	4.161	25.248
4	0.111	0.433	1.109	3.564	23.557
5	0.111	0.430	1.050	3.163	22.043
6	0.111	0.429	1.023	2.895	20.688
7	0.111	0.429	1.011	2.713	19.474
8	0.111	0.429	1.005	2.591	18.386
9	0.111	0.429	1.002	2.508	17.412
10	0.111	0.429	1.001	2.452	16.540
∞	0.111	0.429	1.000	2.333	9.000

(4) $C_s^2=8$

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.150	0.879	2.750	8.050	37.350
2	0.114	0.528	1.747	6.125	34.378
3	0.111	0.452	1.329	4.869	31.725
4	0.111	0.434	1.148	4.034	29.350
5	0.111	0.430	1.067	3.476	27.226
6	0.111	0.429	1.031	3.103	25.324
7	0.111	0.429	1.014	2.852	23.621
8	0.111	0.429	1.007	2.684	22.097
9	0.111	0.429	1.003	2.570	20.732
10	0.111	0.429	1.001	2.494	19.510
∞	0.111	0.429	1.000	2.333	9.000

(5) $C_s^2=10$

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.161	1.007	3.250	9.683	45.450
2	0.114	0.554	1.952	7.190	41.616
3	0.111	0.458	1.416	5.572	38.197
4	0.111	0.436	1.185	4.499	35.137
5	0.111	0.430	1.083	3.784	32.400
6	0.111	0.429	1.038	3.307	29.951
7	0.111	0.429	1.017	2.988	27.759
8	0.111	0.429	1.008	2.774	25.798
9	0.111	0.429	1.004	2.631	24.042
10	0.111	0.429	1.002	2.534	22.470
∞	0.111	0.429	1.000	2.333	9.000

クになることをある程度防ぐことができる．このようなバイパス効果は ρ が余り大きくない場合に顕著である．なぜならば，ρ が大きくなると，極端に大きなサービス要求量を持つジョブが複数個同時にシステムに存在する確率が大きくなり，後続のジョブがバイパスしにくくなると考えられるからである．表1はまた，m-LiPS 規律と PS 規律 (∞-LiPS) の違いをはっきりと示している．変動係数が1に近く，ρ がそれほど大きくない場合は二つの規律の間にそれほどの差異は見られないが，変動係数が大きくなるにつれ，また ρ が1に近づくにつれ，多重度が有限であることの影響が大きくなる．現実には，多くのジョブのサービス要求量のばらつきは大きく，分布の変動係数は1より大きいとされているので，性能評価モデルの中に PS 規律を取り入れる場合は，混雑さ加減をかなり過小評価している，ということに注意しなければならない．

サービス要求量の変動係数が1より小さい場合の例として，アーラン分布を取り上げる．フェーズ k のアーラン分布（これを E_k と表す）の変動係数は $1/k$ である．この場合も超指数分布の場合と同様に陽解が得られていないので，平衡方程式をたて計算機によって $L(m)$ を計算することになる．ここで用いたアルゴリズムでは $\binom{m+k-1}{m}$ 次元の行列2次方程式を解かなければならないので，計算時間の制約から余り大きなシステムは解くことができない．いくつかの計算例を表2にまとめる．この表によれば，変動係数が1より小さい場合は変動係数が1より大きい表1の場合と全く対照的な結果が見て取れる．すなわち，変動係数が1より遠ざかるにつれ，また ρ が1に近づくにつれ，多重処理がシステム効率を悪化させていることが分かる．一方 PS 規律との比較に関して言えば，有限多重度の多重処理を PS 規律でモデル化することは，混雑さ加減を過大評価することになる．

サービス要求量が一般の場合には厳密解を求めることができないので，平均システム内ジョブ数を定量的に調べるために近似式が必要となる．ここで，前章で提案された近似式の精度を，検討してみよう．

サービス要求量が超指数分布の場合，多重度が2ならば，式(5.3)より，

表2 M/E$_k$/1(m-LiPS) における平均システム内ジョブ数
Table 2 Mean queue length $L(m)$ in M/E$_k$/1(m-LiPS).

(1) M/E$_2$/1(m-LiPS)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.108	0.396	0.875	1.925	6.975
2	0.111	0.419	0.938	2.048	7.178
3	0.111	0.426	0.969	2.133	7.360
4	0.111	0.428	0.984	2.193	7.524
5	0.111	0.428	0.992	2.235	7.671
6	0.111	0.429	0.996	2.264	7.804
7	0.111	0.429	0.998	2.285	7.923
8	0.111	0.429	0.999	2.299	8.030
∞	0.111	0.429	1.000	2.333	9.000

(2) M/E$_3$/1(m-LiPS)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.107	0.386	0.833	1.789	6.300
2	0.111	0.416	0.916	1.951	6.569
3	0.111	0.425	0.958	2.065	6.812
4	0.111	0.427	0.979	2.145	7.030
5	0.111	0.428	0.989	2.201	7.226
6	0.111	0.429	0.995	2.240	7.403
7	0.111	0.429	0.997	2.268	7.562
8	0.111	0.429	0.999	2.287	7.705
∞	0.111	0.429	1.000	2.333	9.000

(3) $M/E_4/1$(m-LiPS)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.107	0.380	0.813	1.721	5.963
2	0.111	0.414	0.905	1.903	6.265
3	0.111	0.424	0.952	2.030	6.537
4	0.111	0.427	0.976	2.120	6.782
5	0.111	0.428	0.988	2.183	7.003
6	0.111	0.428	0.994	2.228	7.202
∞	0.111	0.429	1.000	2.333	9.000

(4) $M/E_5/1$(m-LiPS)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.107	0.377	0.800	1.680	5.760
2	0.111	0.413	0.898	1.873	6.082
3	0.111	0.424	0.948	2.009	6.373
4	0.111	0.427	0.974	2.105	6.634
5	0.111	0.428	0.987	2.173	6.869
∞	0.111	0.429	1.000	2.333	9.000

(5) $M/E_6/1$(m-LiPS)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.107	0.375	0.792	1.653	5.625
2	0.111	0.412	0.894	1.854	5.961
3	0.111	0.423	0.946	1.995	6.263
∞	0.111	0.429	1.000	2.333	9.000

$$L(2)-L_a(2)=(K-1)\rho(R-\rho) \tag{5.4}$$

となるから，正確に誤差を計算することができる．$R-\rho$ は，ρ の全域で 0.13 を越えないので，K が大きくないとき $L_a(2)$ は良い近似を与える．多重度が 3 以上の場合は表 1 を使って相対誤差を計算する．**表 3** はこれらの結果をまとめたものである．この表によれば，$C_s^2=2$ の時，相対誤差は 1% 以下であり，$L(m)$ の値を 2 桁しか必要としないならば近似式で十分，ということが分かる．また，$C_s^2=10$ の場合最大誤差でもほぼ 10%，多くの場合 5% 以下で押さえられているので十分実用的であると思われる．さらにこの表から分かることは，近似式が厳密解より常に大きいということである．これは近似式が混雑さ加減を実際に比べて安全側に評価していることを意味する．

サービス要求量がアーラン分布の場合，表 2 の範囲で近似式の相対誤差は最大でも 0.2% を越えない．これは近似式と厳密解がほぼ 3 桁まで一致しているということであるから，実用的には十分過ぎるほどの精度を保証していると言える．

表 3 $M/H_2/1$(m-LiPS) における近似式 $L_a(m)$ の精度
Table 3 Relative error of the approximation formula $L_a(m)$ in $M/H_2/1$(m-LiPS).

(1) $C_s^2=2$ (%)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.0	0.0	0.0	0.0	0.0
2	0.1	0.8	1.0	0.6	0.1
3	0.1	0.4	0.9	0.8	0.1
4	0.0	0.1	0.6	0.8	0.2
5	0.0	0.0	0.3	0.7	0.2
6	0.0	0.0	0.2	0.6	0.3
7	0.0	0.0	0.1	0.5	0.3
8	0.0	0.0	0.1	0.4	0.3
9	0.0	0.0	0.0	0.3	0.3
10	0.0	0.0	0.0	0.2	0.3

(2) $C_s^2=10$ (%)

m	ρ 0.1	0.3	0.5	0.7	0.9
1	0.0	0.0	0.0	0.0	0.0
2	1.8	8.7	8.9	4.0	0.5
3	0.3	5.0	10.4	6.5	0.9
4	0.1	2.0	8.1	7.9	1.2
5	0.0	0.7	5.3	8.3	1.6
6	0.0	0.2	3.2	7.9	1.9
7	0.0	0.1	1.8	7.0	2.2
8	0.0	0.0	1.0	5.9	2.5
9	0.0	0.0	0.5	4.8	2.7
10	0.0	0.0	0.3	3.8	2.9

注) 各数字は $100\frac{L_a(m)-L(m)}{L(m)}$ を表す．

表 1，表 2 によると，$L(m)$ の単調性，すなわちサービス要求量の分布が NBUE (NWUE) なら，$L(m) \geqq L(m-1)$ $(\leqq)$ が成立している．この事実の理論的な証明は現時点ではできなかった．これを今後の研究課題としたい．

謝辞 本研究は日本電気(株)の援助を受けて行った．

本研究に関して多くの貴重な助言をいただいた，川島武先生（防衛大），紀一誠氏（日本電気 C&C システム研究所），宮沢政清先生（東京理科大），およびレフェリーに感謝いたします．

参 考 文 献

1) Brumelle, S.: Some Inequalities for Parallel-Server Queues, *Oper. Res.*, Vol. 19, pp. 402–412 (1971).

2) Doob, J. L.: *Stochastic Processes*, p. 558,

Wiley, New York (1953).
3) Miyazawa, M.: The Derivation of Invariance Relations in Complex Queueing Systems with Stationary Inputs, *Adv. Appl. Prob.*, Vol. 15, pp. 874-885 (1983).
4) Stidham, S.: On the Optimality of Single Server Queuing Systems, *Oper. Res.*, Vol. 18, pp. 708-732 (1970).
5) Stoyan, D.: *Comparison Methods for Queues and Other Stochastic Models*, p. 20, John Wiley & Sons, New York (1983).
6) Yamazaki, G. and Sakasegawa, H.: An Optimal Design Problem of Limited Processor Sharing Systems, Institute of Socio-Economic Planning, Discussion Paper Series, No. 246, University of Tsukuba (1984), submitted to Management Science.
7) 山崎源治：Erlang's loss formula について，京都大学数理解析研究所講究録，No. 452, pp. 262-272 (1982).

(昭和 60年6月 27日受付)
(昭和 61年8月 27日採録)

山崎　源治

昭和21年生．昭和43年工学院大学機械工学科卒業．昭和50年同大学大学院工学研究科博士課程満期退学．同年同大学生産機械工学科助手．昭和61年東京都立科学技術大学管理工学科助教授．工学博士．確率過程とその応用（主に，待ち行列・生産システム）に興味を持つ．日本 OR 学会，日本機械学会，日本経営工学会各会員．

逆瀬川浩孝（正会員）

1944年生．1969年東京大学理学部数学科卒業．同年文部省統計数理研究所入所．1980年筑波大学社会工学系助教授．待ち行列モデルの数値解析，シミュレーションの方法論，乱数の生成法などの研究に従事．日本オペレーションズリサーチ学会，日本統計学会，応用統計学会各会員．