集団到着のある無限サーバ待ち行列とその関連モデル

矢島　萌子

無限個のサーバをもつ待ち行列モデルである無限サーバ待ち行列は，システムに到着した客は到着後すぐにサービスを受け始めることができる「待ち行列」が発生しない待ち行列モデルです．無限サーバ待ち行列は，大規模施設の系内客数の挙動・交通システム・在庫システムなど，さまざまな状況をモデル化する際に使用されます．本稿では，客が集団で到着する無限サーバ待ち行列に関する研究を，著者が執筆した博士論文の内容をもとに紹介します．

キーワード：待ち行列理論，無限サーバ待ち行列，集団到着型待ち行列

1.　はじめに

無限サーバ待ち行列とは，無限個のサーバをもつ待ち行列です．無限サーバ待ち行列は，システムに到着した客は到着後すぐにサービスを受け始めることができるため，「待ち行列」が発生しない待ち行列モデルです．私は，無限サーバ待ち行列のなかで，特に客が集団で到着するモデルに興味をもち研究を行ってきました．

現実に無数のサービス機能が存在する状況というのは想像しにくいため，無限サーバ待ち行列の応用例はイメージしにくいかもしれません．しかし，実際は無限サーバ待ち行列はさまざまな状況をモデル化するために使用できます．たとえば，大規模施設の客数の挙動・交通システム・在庫システムなどをモデル化できます．また，無限サーバ待ち行列は有限サーバ待ち行列を近似的にモデル化する際にも使われます．

本稿では，私が行ってきた，集団到着のある無限サーバ待ち行列に関する研究について，博士論文の内容をもとにして紹介します．2 節では，集団到着のある無限サーバ待ち行列の安定性についてお話しします．次に，背後過程をもつ集団到着のある無限サーバ待ち行列の極限解析について，3 節で紹介します．4 節では，無限サーバ待ち行列の関連モデルとして，処理速度が可変な単一サーバ待ち行列の解析についてお話しします．

2.　無限サーバ待ち行列の安定性

安定な待ち行列は，長時間にわたって稼働してもシステム内の客数が（確率 1 で）発散しません．待ち行列の安定性は，システムが安全に稼働するための重要な条件といえます．数学的には，客数過程が proper かつ非退化な極限分布をもつときその待ち行列が安定であると定義されます．また，待ち行列が安定であるための必要十分条件を安定条件と呼びます．

有限サーバ待ち行列の安定条件に関する Loynes [1] の定理は，待ち行列理論の最も有名な定理の一つです．Loynes [1] の定理は，客が集団で到着する場合も適用できますが，無限サーバ待ち行列には適用できません．客が 1 人ずつ到着する無限サーバ待ち行列は，平均到着間隔と平均サービス時間の有限性が安定条件です [2]．

集団到着のある場合，無限サーバ待ち行列の安定条件はどうなるでしょう．実は，集団到着のある無限サーバ待ち行列の先行研究では，適当な十分条件を仮定したり，定常分布の存在そのものを（暗に）仮定して議論をしていて，安定性の保証のために最低限必要な条件を仮定して解析を進めているものはそれほど多くはありません．ポアソン到着・指数サービス時間の基本的な集団到着のある無限サーバ待ち行列である $\mathrm{M}^{\mathrm{X}}/\mathrm{M}/\infty$ 待ち行列については，Cong [3] が安定条件は集団サイズの対数期待値の有限性であると示しています．集団サイズとは，一つの集団に属する客の数のことです．Cong [3] の結果から，集団到着のある無限サーバ待ち行列の安定条件が，到着率・集団サイズ・サービス率の平均値だけでは議論できないことがわかります．

2.1　BMAP/M/∞ 待ち行列の安定性

本節では，$\mathrm{M}^{\mathrm{X}}/\mathrm{M}/\infty$ 待ち行列よりも一般的なモデルである $\mathrm{BMAP}/\mathrm{M}/\infty$ 待ち行列の安定条件を紹介します．$\mathrm{BMAP}/\mathrm{M}/\infty$ 待ち行列は，BMAP (Batch Markovian Arrival Process [4]) と呼ばれる集団到着

やじま　もえこ
東京工業大学情報理工学院数理・計算科学系
〒 152–8552 東京都目黒区大岡山 2–12–1

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

過程に従って集団が到着する無限サーバ待ち行列です．客のサービス時間は指数分布に従います．

BMAP は，有界な状態空間 $\mathbb{D}$ 上の既約な連続時間マルコフ連鎖 $\{J(t)\}$ に依存して到着率（強度）や到着する集団サイズの分布が変化する集団到着過程です．BMAP は行列の集合 $\{\boldsymbol{D}(k) = (D_{i,j}(k))_{i,j\in\mathbb{D}}; k \in \mathbb{Z}_+\}$ をパラメータとしてもちます．$\mathbb{Z}_+$ は非負整数の集合です．$N(t)$ を時刻 $[0, t)$ の間に到着した客の総数とすると，パラメータ $\{\boldsymbol{D}(k); k \in \mathbb{Z}_+\}$ で特徴づけられる BMAP は，次式で定義されます．

$$\mathsf{P}(N(t+\Delta t) - N(t) = k, J(t+\Delta t) = j | J(t) = i) = \begin{cases} 1 + D_{i,i}(0)\Delta t + o(\Delta t) & k = 0, i \in \mathbb{D}, \\ D_{i,j}(k)\Delta t + o(\Delta t) & \text{その他}. \end{cases}$$

$\boldsymbol{D}(0)$ は負の対角成分と非負の非対角成分をもつ優対角行列，$k = 1, 2, \ldots$ で $\boldsymbol{D}(k)$ は非負行列です．また，$\sum_{k=0}^{\infty} \boldsymbol{D}(k)\boldsymbol{e} = \boldsymbol{0}$ を満たします．$\boldsymbol{e}$ と $\boldsymbol{0}$ はすべての成分が，それぞれ 1 と 0 である列ベクトルです．

BMAP は集団ポアソン到着過程など，多くの集団到着過程を包含しています．また，BMAP はパラメータを適切に設定することで，任意の集団到着過程を任意の精度で近似できることが知られています．このことから，BMAP はとても一般的な集団到着過程といえます．そして，このように定義される BMAP/M/∞ 待ち行列の安定条件は，以下のように与えられます．

定理 1. BMAP/M/∞ 待ち行列の安定条件 [5]
BMAP/M/∞ 待ち行列の安定条件は，集団サイズの対数期待値 $\mathsf{E}[\log(X)]$ が有限であること

ある確率変数 Z の対数期待値は，$\mathsf{E}[\log(Z)]$ のことを指します．よって，BMAP/M/∞ 待ち行列の安定条件は，BMAP のパラメータを用いて

$$\sum_{k=1}^{\infty} \log k \cdot \boldsymbol{D}(k)\boldsymbol{e} < \infty,$$

と書けます．定理 1 の結果は，集団サイズの期待値が発散してもシステムが安定になり得ることを意味しています．この事実は直観的に意外かもしれません．

2.2 GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定性

2.1 節では，サービス時間が指数分布に従うモデルの安定性について考えました．次に，サービス時間が一般的な分布に従う集団到着のある無限サーバ待ち行列の安定性に関する研究を紹介します．

GI$^{\mathrm{X}}$/GI/∞ 待ち行列は，集団が再生過程に従って到着して，集団サイズが確率変数 X に従う無限サーバ待ち行列です．客のサービス時間は独立に一般的な分布関数 $H(x)$ に従います．$S_1, S_2, \ldots$ を互いに独立に分布 $H(x)$ に従う確率変数列とすると，GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件は以下のように与えられます．

定理 2. GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件 [2, 6]
GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件は，次の不等式が成立すること

$$\mathsf{E}\left[\max_{k=1,\ldots,X} S_k\right] < \infty.$$

Pakes and Kaplan [2] は，GI$^{\mathrm{X}}$/GI/∞ 待ち行列の客数過程が Bellman-Harris 過程の特別な場合であることに着目して，Bellman-Harris 過程の極限分布の存在条件として定理 2 を示しています．Bellman-Harris 過程は，年齢依存の分岐過程の一つで生物集団の個体数の時間的変化を表すために用いられます．一方で，Yajima and Masuyama [6] はこの定理に対して待ち行列理論的なアプローチで別証明を与えています．

定理 2 の期待値の中身は，同一集団内のサービス時間の最大値です．この期待値は集団サイズ X とサービス時間 $S_1, S_2, \ldots$，2 種類の一般的な分布に従う確率変数が入っていて，計算しにくいです．期待値が計算しにくいと不等式が成立しているか（安定条件を満たしているか）の判定が難しく，実用上で使いづらい安定条件となります．そこで，GI$^{\mathrm{X}}$/GI/∞ 待ち行列に何か条件を加えることで，扱いやすい安定条件を導出できる場合がないか考えます．

サービス時間分布 H の裾が指数分布と同じようにふるまう場合，つまり裾が指数的に減衰する場合を考えます．分布 H の裾が指数的に減衰するとは，次の不等式を満たす $\alpha, \beta > 0$ が存在することです．

$$0 < \liminf_{x\to\infty} \frac{1 - H(x)}{\mathrm{e}^{-\alpha x}} < \limsup_{x\to\infty} \frac{1 - H(x)}{\mathrm{e}^{-\beta x}} < \infty.$$

サービス時間分布の裾が指数的に減衰するとき，GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件は以下のように与えられます．

系 1. GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件 [6]
サービス時間分布の裾が指数的に減衰するとき GI$^{\mathrm{X}}$/GI/∞ 待ち行列の安定条件は，集団サイズの対数期待値が有限であること

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

$\mathrm{GI^X/GI/\infty}$ 待ち行列では，サービス時間分布は裾さえ指数的に減衰すれば指数分布でなくても，$\mathrm{BMAP/M/\infty}$ 待ち行列など指数サービス時間のモデルと同様の安定条件をもつことがわかります．

3. 背後過程をもつ無限サーバ待ち行列

待ち行列システムの背後に存在する連続時間マルコフ連鎖に依存してパラメータが変化する待ち行列モデルは，マルコフ変調待ち行列と呼ばれます．背後にあるこのマルコフ連鎖は背後過程と呼ばれ，その推移は待ち行列システムの状態に依存しません．

マルコフ変調待ち行列は複雑なシステムの確率的なふるまいを表現できますが，モデルの複雑さゆえに陽な形で各指標（系内客数など）を導出することは難しいです．背後過程が 2 状態の場合などモデルの設定が簡潔である場合には，陽解析を行っている先行研究も存在します．しかし，マルコフ変調待ち行列を扱う多くの先行研究では，極限解析を行うことでモデルの特性を把握しています．

3.1 カタストロフィ現象の発生を伴うマルコフ変調無限サーバ待ち行列の解析

本節では，マルコフ変調 $\mathrm{M^X/M/\infty}$ 待ち行列の極限解析を紹介します．$\mathrm{M^X/M/\infty}$ 待ち行列とは，ポアソン到着・指数サービス時間の集団到着のある無限サーバ待ち行列です．今回は，このマルコフ変調 $\mathrm{M^X/M/\infty}$ 待ち行列にカタストロフィ現象と呼ばれるイベントの発生が伴うモデルを考えます．カタストロフィ現象を考えることで，外的要因によって（サービスが完了していなくても）客が退去する可能性のある状況をモデル化することができます．

カタストロフィ現象とは，パソコンのシステム障害・ウイルス感染・人口モデルにおける自然災害など，現実世界におけるアクシデントを表現します．カタストロフィ現象が発生すると，その時点で待ち行列システム内に滞在する客は，ある確率でサービスを完了することなくシステムから退去します．今回は，カタストロフィ現象が定常ポアソン過程に従って発生して，発生したときにシステム内にいる客はそれぞれ独立に確率 p でシステムに残り，確率 $1-p$ でシステムから退去する場合を考えます．

極限解析を行うために，スケーリングモデルを構築します．スケーリングパラメータとして，ファクター N と係数 α を用意します．スケーリングモデルでは，到着率を N 倍，背後過程の推移率を N^α 倍します．ファクター N を無限に大きくすると，到着率は発散するのでスケーリングモデルの重負荷極限を扱うことになります．さらに，カタストロフィの発生率を N 倍，発生時の客の退去確率を N^{-1} 倍します．

$L^{(N)}$ を到着率が N 倍されたスケーリングモデルの定常系内客数を表す確率変数とします．まずは，以下の補題を紹介します．

補題 1.　定常系内客数の大数の法則 [7]
確率変数 $L^{(N)}/N$ は定数 ρ に確率収束する

確率収束先の定数 ρ は，与えられたシステムパラメータから簡単に計算できる値です．紙面の都合上，ρ の具体的な値の紹介は割愛させていただきますが，ご興味のある読者の方々は文献 [7] をご覧ください．

補題 1 より，$(L^{(N)}/N-\rho)$ は $N\to\infty$ で 0 に確率収束することがわかります．それでは，β をある実数として，$N\to\infty$ としたときに $N^\beta(L^{(N)}/N-\rho)$ はどのようなふるまいをみせるでしょうか．$\beta=\max(1,\alpha)/2$ のとき，$N^\beta(L^{(N)}/N-\rho)$ は，以下のとおり発散することも 0 に収束することもありません．

定理 3.　定常系内客数の中心極限定理 [7]
$\beta=\max(1,\alpha)/2$ とすると，確率変数

$$N^\beta\left(\frac{L^{(N)}}{N}-\rho\right)$$

は平均 0，分散 σ^2 の正規分布に弱収束する

定数 σ^2 についても ρ と同様に与えられたシステムパラメータから簡単に計算できます．無限サーバ待ち行列は各サーバが独立にふるまうこと，ポアソン過程の和はポアソン過程になることから，定理 3 の内容は直観的にイメージしやすいと思います．ただし，$N^\beta(L^{(N)}/N-\rho)$ が 0 でない値に収束するための係数 β が係数 α に依存するのは，少し意外かもしれません．

Yajima and Phung-Duc [7] は，$N^\beta(L^{(N)}/N-\rho)$ の特性関数が正規分布の特性関数に各点収束することを示すことで定理 3 を証明しました．Lévy の連続定理より特性関数の各点収束と分布の弱収束は同値なので，この方法で証明することができます．この特性関数を用いた中心極限定理の証明方法は古典的な方法ですが，分布の収束速度について議論できないという欠点があります．

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

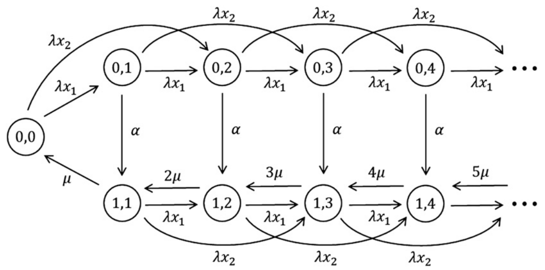

図 1 $x_1+x_2=1$ の場合のマルコフ連鎖 $\{(J(t),L(t))\}$ の状態遷移図
ただし，$x_1=\mathsf{P}(X=1)$，$x_2=\mathsf{P}(X=2)$.

4. 処理速度が可変なサーバをもつモデル

処理速度が可変なサーバをもつ待ち行列モデルの研究は，コンピュータシステムにおけるエネルギー消費増加と遅延時間増大の問題が大きな動機となっています．処理速度を速くするとサーバの運用コスト（エネルギー消費）が大きくなりますが，エネルギー消費を気にして速度を遅くしすぎるとシステムが混雑してユーザーの満足度を低下させます．エネルギー消費削減と遅延時間短縮はトレードオフの関係にあります．そのため，処理速度が可変なサーバをもつ待ち行列モデルを解析することは，ユーザーが我慢できるくらいに遅延時間を抑えながらエネルギー消費を削減するような処理速度の変更基準を知ることに繋がります．

また，待ち行列システムのエネルギー消費を抑えるために，使わないサーバの電源を切るという方法が考えられます．ただし通常は，サーバの電源を入れた後は，客を処理できるようになるまでに起動時間が必要です．起動中はエネルギーを消費しますが，客を処理することはできません．電源のオン・オフの切替が可能な待ち行列モデルにおいて，起動時間が必要なこと・起動中はエネルギー消費をすることは，コンピュータシステムへの応用などを考えれば必要な仮定です．

4.1 系内客数に比例して処理速度を変えるモデル

本節では，系内客数に比例して処理速度を変えるモデルについて詳しく紹介します．集団到着の単一サーバ待ち行列で，集団が率 λ のポアソン過程に従って到着して，集団サイズは確率変数 X に従います．客のサービス要求量は平均 1 の指数分布に従い，系内客数 n のときにサーバは単位時間あたり $n\mu$ のサービスを処理します．システム内の客が空になったらすぐにサーバの電源を切り，空のシステムに再び客が到着したらすぐにサーバの電源を入れます．電源を入れてから客を処理できるまでに，平均 $1/\alpha$ の指数分布に従う起動時間が必要です．

$L(t)$ を時刻 t における系内客数，$J(t)$ をサーバの状態とおきます．$J(t)$ はサーバが処理しているときは 1，していないときは 0 の値をとるとします．このとき確率過程 $\{(J(t),L(t))\}$ は連続時間マルコフ連鎖であり，その状態遷移図は図 1 のようになります．つまり，系内客数に比例して処理速度を変えるこのモデルは，マルコフ解析の枠組みで扱うことができます．

図 1 を見ると，この単一サーバ待ち行列の客数過程が無限サーバ待ち行列の客数過程と等価であることがわかります．マルコフ連鎖 $\{(J(t),L(t))\}$ の状態 (i,j) における定常分布を $\pi_{i,j}$ とおきます．さらに，サーバの状態が 0 のときの母関数を $\widehat{\pi}_0(z)=\sum_{j=0}^{\infty}\pi_{0,j}z^j$，サーバの状態が 1 のときの母関数を $\widehat{\pi}_1(z)=\sum_{j=1}^{\infty}\pi_{1,j}z^j$ とします．このモデルの定常系内客数の確率母関数は，以下のように陽な形で表されます．

定理 4. 定常系内客数分布 [8]
定常系内客数の確率母関数 $\widehat{\pi}(z)$ は次式で与えられる．

$$\begin{aligned}\widehat{\pi}(z)&=\widehat{\pi}_0(z)+\widehat{\pi}_1(z),\\ \widehat{\pi}_0(z)&=\frac{\lambda+\alpha}{\lambda+\alpha-\lambda\widehat{X}(z)}\pi_{0,0},\\ \widehat{\pi}_1(z)&=\int_0^z \mathrm{e}^{\frac{\lambda}{\mu}(Q(z)-Q(u))}\frac{\lambda}{\mu}q(u)\widehat{\pi}_0(u)\mathrm{d}u.\end{aligned}$$

ただし，$\widehat{X}(z)=\mathsf{E}[z^X]$，$q(z)=(1-\widehat{X}(z))/(1-z)$，$Q(z)=\int_0^z q(u)\mathrm{d}u$，$\pi_{0,0}$ は $\widehat{\pi}(1)=1$ を満たす値

マルコフ連鎖 $\{(J(t),L(t))\}$ の平衡方程式から $\widehat{\pi}(z)$ に関する微分方程式を導き，その微分方程式を解くことで $\widehat{\pi}(z)$ を得ることができます．このようにマルコフ連鎖の平衡方程式を利用して確率母関数を導く解析手法は母関数法と呼ばれ，マルコフ連鎖の定常分布を知るためによく用いられます．

定理 4 の結果を用いて，エネルギー消費に関する考察を行うことができます．系内客数 n のときに速度 $n\mu$ で処理中のサーバは単位時間あたり $(n\mu)^2$ でエネルギーを消費するとします．また，起動中のサーバは単位時間あたり μ^2 でエネルギーを消費するとします．電源がオフのときのサーバはエネルギーを一切消費しません．このとき，単位時間あたりの平均消費エネルギー $\mathcal{E}$ は定理 4 で導出した $\widehat{\pi}(z)$ を用いて，以下のように書き表すことができます．

$$\mathcal{E}=\mu^2\cdot\widehat{\pi}_1(1)''+\mu^2\cdot(\widehat{\pi}_0(1)-\pi_{0,0}).$$

さらに，Yajima and Phung-Duc [8] は，このモデ

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.

ルの滞在時間分布についても解析を行い，定常滞在時間のラプラス変換を導出しています．このモデルでは，系内客数によって処理速度が変化するため，ある客の滞在時間はその客より後に到着する客の影響を受けてしまいます．そのため，ある客の滞在時間が到着時のシステムの状態だけでは定まらず，滞在時間に関する解析が難しくなっています．今回は系内客数に比例させるという比較的単純な方法で処理速度を変化させていたため滞在時間のラプラス変換を得ることができましたが，もう少し複雑な基準で処理速度を変化させるモデルでは滞在時間分布の陽解析はさらに難しいです．

Yajima and Phung-Duc [9] は，サーバのオン・オフ規律に関する設定はこのモデルと同じで，ほかの基準で処理速度を変更するようなモデルを扱っています．また，サーベイ論文 [10] では，処理スピードが可変である待ち行列モデルに関する多くの研究が紹介されているので興味のある読者は目を通してみてください．

5. おわりに

本稿では，私のこれまで行ってきた研究について博士論文の内容を中心に紹介させていただきました．さまざまな内容を一緒くたに詰め込んでしまったため，機関誌の記事としては読みにくかったかもしれません．

苦慮することもありましたが，私は多くの方々に支えられて，三好直人教授（東京工業大学）のもとで博士号を取得しました．そして，2020 年 4 月から東京工業大学の情報理工学院に助教として，研究者人生をスタートさせています．読者のなかには，博士課程に在籍中，進学予定，もしくは進学を検討中の学生さんもいらっしゃるかもしれません．そのような学生さんが博士論文執筆までのイメージを思い浮かべるお手伝いを本稿ができていれば幸いです．

参考文献

[1] R. M. Loynes, “The stability of a queue with non-independent inter-arrival and service times,” *Mathematical Proceedings of the Cambridge Philosophical Society*, **58**, pp. 497–520, 1962.

[2] A. G. Pakes and N. Kaplan, “On the subcritical Bellman-Harris process with immigration,” *Journal of Applied Probability*, **11**, pp. 652–668, 1974.

[3] T. D. Cong, “On the $M^X/G/\infty$ queue with heterogeneous customers in a batch,” *Journal of Applied Probability*, **31**, pp. 280–286, 1994.

[4] D. M. Lucantoni, “New results on the single server queue with a batch Markovian arrival process,” *Stochastic Models*, **7**, pp. 1–46, 1991.

[5] M. Yajima, T. Phung-Duc and H. Masuyama, “The stability condition of BMAP/M/∞ queues,” In *Proceedings of the 11th International Conference on Queueing Theory and Network Applications (QTNA2016)*, Article number 5, 2016.

[6] M. Yajima and H. Masuyama, “Stability analysis of $GI^X/GI/\infty$ queues,” In *Proceedings of the 14th International Conference on Queueing Theory and Network Applications (QTNA2019)*, pp. 27–29, 2019.

[7] M. Yajima and T. Phung-Duc, “A central limit theorem for a Markov-modulated infinite-server queue with batch Poisson arrivals and binomial catastrophes,” *Performance Evaluation*, **129**, pp. 2–14, 2017.

[8] M. Yajima and T. Phung-Duc, “Batch arrival single-server queue with variable service speed and setup time,” *Queueing Systems*, **86**, pp. 241–260, 2017.

[9] M. Yajima and T. Phung-Duc, “Analysis of a variable service speed single server queue with batch arrivals and general setup time,” *Performance Evaluation*, **138**, pp. 241–260, 2020.

[10] M. Yajima and T. Phung-Duc, “Queues with variable service speeds: Exact results and scaling limits,” *The Palgrave Handbook of Operations Research*, S. Salhi and J. Boyan (eds.), Springer (to appear).

 Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.