待ち行列分野のオープンプロブレム

高木　英明，大野　勝久，宮沢　政清，町原　文明，紀　一誠

現在，情報化は目覚ましい勢いで進んでおり，待ち行列に関する研究も変化している．様々な分野で解決すべき新しい課題が多様な形で現れてきている．これらの課題についていろいろな角度から捉え，解決し豊かな研究成果をあげ社会に貢献していくことが望まれている．本稿では，待ち行列理論の無線通信網とサービスおよび生産システムへの応用，待ち行列理論の研究課題，待ち行列理論の深奥とORの関係，参照局所性の数理的基礎付けについてオープンプロブレムを提起している．

キーワード：待ち行列理論，$GI/G/\infty$，参照局所性，サービス，生産システム，無線通信網

1.　はじめに

待ち行列研究部会　主査 小野里好邦
幹事 河西　憲一

現在，情報化は目覚ましい勢いで進んでいる．情報処理技術の急速な発展，普及により今まで経験したことがない変革がおこり，産業，社会生活も変わってきている．日本オペレーションズ・リサーチ学会が対象とする研究も変化を遂げている．それに伴い，様々な分野で解決すべき新しい課題が多様な形で現れてきている．将来を見通しこれらの課題について様々な角度から前向きに捉え，解決し豊かな研究成果をあげ社会に貢献していくことが望まれている．

待ち行列研究部会は，2007年7月の定例研究部会で200回を数えた．これまで待ち行列研究部会が対象にしていた研究内容に変化の兆しが現れてきている．待ち行列に関しても，社会が変化していく中で，新しい多様な側面を呈し，解決すべき多数の問題が現れている．これからの情報化社会の行く末を見通し，待ち行列という枠組みを大幅に緩めて，さらにはその枠組みを超えて解決検討すべき課題について，その手法，解決手段，研究方法を議論することは将来の社会が抱える問題を解決する上で極めて重要である．

本稿は，待ち行列分野には大きな研究領域が広がっていることを十分に意識している研究者による各分野における問題意識とその概念，将来の研究に関する展望についてまとめたものである．具体的には，待ち行列理論の応用分野として無線通信網とサービスおよび生産システムに関連したオープンプロブレム，待ち行列理論の研究課題，待ち行列理論の深奥とORの関係，参照局所性の数理的基礎付けについて問題を提起している．

2.　待ち行列理論の応用への期待

高木　英明

1970代以降の待ち行列理論の進歩は，数学的動機によるものも多いが，計算機や情報通信網の性能評価への応用に牽引された点が大きいことは否めない．顕著な例として，計算機のオペレーティングシステムをモデル化した待ち行列網の積形式解（BCMP定理），トークンリング方式のローカルエリアネットワークのための巡回型サービス（ポーリング）モデルとその単純形であるサーバ・バケーションモデル，広帯域通信網のATM交換機におけるセル到着過程をもつ待ち行列のマトリックス幾何解法等が挙げられる．しかしながら，計算機や通信プロトコルの制御が強化されるにつれ，関連する動作過程の独立性に依拠する待ち行列理論は適用が難しくなった．また，CPUの性能向上や光ファイバーの通信帯域の爆発的増加により，そ

たかぎ　ひであき
筑波大学 大学院システム情報工学研究科
〒305-8573 つくば市天王台1-1-1
おおの　かつひさ
愛知工業大学 経営情報科学部
〒470-0392 豊田市八草町八千草1247
みやざわ　まさきよ
東京理科大学 理工学部
〒278-8510 野田市山崎2641
まちはら　ふみあき
東京電機大学 理工学部
〒350-0394 埼玉県比企郡鳩山町
きの　いっせい
神奈川大学 理学部
〒259-1293 平塚市土屋2946

れらは「待ち」を要するボトルネックではなくなってしまった．そのような状況の中で最近の研究を見ると，待ち行列の理論モデルの解析はますます精巧になる一方で，現実のシステムの性能評価はシミュレーションのみに頼り，両者の乖離が著しい．

本稿では，待ち行列理論が再び応用に貢献できると思われる無線通信網とサービスについて，我田引水の誹りは覚悟の上で筆者の関わりを含めて記し，現実の技術への貢献を通しての待ち行列理論の発展を期待したい．

筆者らは，モバイルITフォーラムの要請により，移動体通信網における2010年以降の多様なサービスに対するユーザの品質要求を満たすために必要な無線周波数帯域幅の算出方法を国際電気通信連合無線通信部門（ITU-R）に提案し，世界無線通信会議WRC-2007で採択された．そこで必要とされたのは，回線交換型（呼損系）とパケット交換型（待時系）のそれぞれ複数のサービス・クラスのパケットをクラスごとに指定された呼損率または平均待ち時間というサービス品質を満たすために必要な通信チャネル（サーバ）の数を決める簡単な（すなわち，各国の標準化担当者が理解でき，スプレッドシートに実装できる）計算法であった．そのために，回線交換型サービスには多次元Erlang-B公式に対するKaufman-Robertsの計算法を用いたが，パケット交換型サービスには，複数クラスの客をもつM/G/m待ち行列の平均待ち時間の適当な近似式が見つからなかったので，これを非割り込み型優先処理M/G/1待ち行列に対するCobhamの式で代用した．後者は現実の処理方式を反映していない代替モデルであり，その提案の採択には忸怩たるものがある．将来の社会基盤である情報通信網の設計にこのような（不完全なモデルに基づく）方法を使わざるを得ない事実には，理論研究の意義が問われる．

最近，サービス・セクターの生産性向上とイノベーションが叫ばれる中で，従来は商学で扱われていたサービス業務に数理・統計・情報科学等を適用しようという「サービス科学」が提唱された．筆者は，平成18年度から「サービスの科学：理論と実践」という科目を開講し，文科系修士レベルの学生に，待ち行列理論によるサービス施設のモデルを教えている．サービスのモデル化では，OR分野の研究者が思いつかなかったような様相も多々ありそうである．例えば，筆者の授業で，レストランにおけるサービスは何でサーバは誰（何）であると思うかと学生に聞いたら，サービスは食事をしている時間で，サーバはコックまたはウェイターという答えが多かった．待ち行列研究者の常識では，サービス時間はサーバが稼動している時間である．しかし，客が食事をしているとき，コックはその客の食事を作っているわけではない．待ち行列理論の成果の例としてよく挙げられる「窓口数の効果」も，人を対象とするサービスでは成り立たないことが多い．これは，複数の到着過程を別々の待ち行列に並ばせてそれぞれの窓口でサービスするよりも，待ち行列を1つにまとめて単一の窓口でサービスする方が，客の平均待ち時間が減る（先着順サービスも保証される）ので，銀行の店舗や空港でのチェックインで使われている方式である（待つ場所を節約し，長い行列の印象を薄めるために，客は幾重にも折り返す蛇形に並ばせられる）．しかし，同様の理由で，全国のスターバックスの客を一箇所に集め，数千台のコーヒーメーカーからサービスをすることはできない．なぜなら，顧客は，サービスを受ける間，物理的にその場にいる必要があるので，顧客の移動のために金銭的・時間的コストが伴うからである．

もともとサービス施設のモデルである待ち行列の理論は，サービスが産業や情報技術の中核となる21世紀前半においても新たな貢献の可能性が十分にあり，新しい応用への模索が理論の発展にフィードバックされることを期待している．

3. 生産システムに関連して

大野　勝久

生産システムには，需要変動，設備故障，原材料・部品の供給遅延，作業者の欠勤等々，様々な不確実性が不可避的に付きまとい，基礎理論としての単一工程生産システムの性能評価は，待ち行列理論の独壇場である．しかしながら，現実に直面する多くの問題は，複数の待ち行列がネットワーク状につながった，待ち行列ネットワークの問題である．そこでは，計算機・通信システムはともかく，生産システムにたいしては，「待ち行列理論がここに使われています」とは，なかなか言い難い現状である．実際，多品種・多工程からなる現実の生産システムの性能評価，最適化あるいは最適制御，どれをとってもその厳密な理論解析は，次世代というよりも，むしろ永遠のオープンプロブレムなのかもしれない．それならば一層，積形式解が成り立つ生産システムを構築するのが，次世代のオープンプロブレムであるといって，この稿を終えられれば良

いが，これはまた夢物語にすぎない．

時の経つのは何とやらで，もう一昔前になってしまったが，Buzacott and Shanthikumar[1]は，待ち行列理論を駆使して，単一工程，フローライン，トランスファライン，ジョブショップ，FMS（Flexible Manufacturing System）の性能評価と設計・制御問題を論じている．さらに10章では，受注生産ライン，有限バッファ生産ライン，共有バッファ生産ライン，MRP（Material Requirements Planning），OPT，基点在庫（base stock）方式，かんばん方式，CONWIP等々，従来個別に論じられてきた多くの生産システムを統一的に議論する基盤として，PAC（Production Authorization Cards）システムを提案している．ここでOPT（Optimized Production Technology）は，ゴールドラット（Eliyahu M. Goldratt）博士によって1970年代後半から開発されてきた生産スケジューリングソフトであり，その発展形が，ボトルネック工程等の制約条件を改善して全体最適を図る手法として，一時期ブームとなった制約条件の理論（Theory Of Constraints, TOC）である．PACシステムは，単一品種N工程直列生産ラインにたいして，4種類のかんばん，すなわち，調達タグ（Requisition tag），発注タグ（Order tag），加工タグ（Process tag），生産指示かんばん（Product Authorization card）を導入し，それらを適切に設定することで，上記生産システムを含む10種の生産方式を特別な場合として実現できることを示し，近似的な性能評価手法を与えている．しかし，生産方式の優劣を比較するためには，近似手法では不十分である．さらに，PACシステム後に提案された一般化かんばん方式を含むように，見直しを行うことも必要である．また，多品種に対する原材料・部品の供給から製造，物流，販売に至る関連企業が，企業の枠を越えて一体となり，調達製造物流・販売を通して全体最適を目指すSCM（Supply Chain Management）へと時代が進展している現在，直列生産ラインをSCMネットワークに拡張することも必要である．このように拡張されたPACシステムにたいする厳密な性能評価手法の開発は，次世代ORのオープンプロブレムになるのではないであろうか？　これにより，かんばん方式に代表されるプル方式とMRPに代表されるプッシュ方式との優劣比較という，シミュレーションを用いた多くの研究が既に行われているにもかかわらず，MRPが依然として大勢を占める問題に決着をつけることが期待される．さらに，これもシミュレーションを用いた多くの研究が既に行われている，プル方式間の比較についても，より説得力のある結論が期待される．

参考文献

[1] J. A. Buzacott and I. G. Shanthikumar, Stochastic Models of Manufacturing Systems, Prentice Hall, NJ, 1993.

4.　待ち行列理論に関連して

宮沢　政清

4.1　理論研究の課題

待ち行列理論はランダムな要因があるサービスシステムの設計と運用のための理論である．システムの性能評価はこの第一歩であり，古くから理論研究の主流であった．例えば，マルコフ過程を用いてモデルを表現し，定常分布を解析的または数値的に求める．Neutsの行列幾何形式解はこのような数値解析法に大きな進展をもたらしたが，解を求める方法の理論研究は限界に達している．

待ち行列本来の目的からは，性能評価量を計算するだけではなく，システムの比較や最適な運用方法などに答えることが重要である．これらは理論的に難しい．従来数値計算やシミュレーションに頼ることが多く，直感的な発想に基づく近似式もよく使われてきた．

この困難な問題に取り組むことが今後の待ち行列理論の大きな課題である．このためには，計算が可能で役立つ量を探すことが重要である．例えば，従来から使われてきた平均や分散は必ずしも扱いやすい量ではない．そこで，解析的に魅力ある量として，

(i) 確率過程の時間と状態のスケール変換（縮尺変更）による極限

(ii) 稀少事象の漸近特性

が注目されている．これら2つの量について説明する．

4.2　極限過程による最適化

(i)はモデルの極限であり，流体モデルや，拡散過程モデルが極限として得られる．安定過程（stable process）やフラクタル過程などもこの種のものである．この極限モデルの主目的は近似モデルを作ることではない．例えば，直感的にモデルを当てはめた流体近似や拡散近似と比較することは意味がない．極限モデルの目的は元のモデルの特性を解明することである．例えば，極限として得られた流体モデルから，元のモデルの安定性を論じることができる．拡散過程モデル

は流体モデルの精密化であり，原理的には，元のモデルについてより多くの情報を持っている．実際に，極限として得られた拡散過程モデルを使って，最適なサービス方法が論じられている．

このような極限モデルの特徴はスケール変換について不変なことである．この不変性を使うとモデルのパラメーターが変化したときにも柔軟に対応できる．また，その柔軟さからマルコフ決定過程などの最適化理論と結びついて，大きな発展が期待される．

4.3 稀少事象の漸近特性

(ii)では，評価量のスケール変換を行って漸近的な特性を求める．例えば，待ち行列長の定常分布の裾の漸近特性がある．この漸近特性を使って，待ち行列ネットワークのサービス方法の比較ができる．

この分野の一般的な方法論に大偏差値理論があり，減少率が最適化問題の解として得られている．しかし，この最適化問題の変数は関数であり，解析的に解を得ることが難しい．特に，2つ以上の待ち行列がある待ち行列ネットワークに対しては，直列型などの簡単な場合を除いて，定常分布の裾の漸近特性を求めることができない．

一般に，n個の待ち行列を持つモデルは反射壁をもつn次元マルコフ変調ランダムウォークとしてモデル化できる．このモデルの定常分布の裾の減少率を，大偏差値理論ではなく，マルコフ再生定理を使って精密に求める研究が進められている．この方法により，2次元までならば，一般的な境界推移の下で減少率を求めることができつつある．例えば，ネットワークにおいて，ノードが空になるとサービス係が別のノードに応援に行くモデルの効果を調べることができる．

これからの大きな課題は3次元以上の問題である．例えば，n本の待ち行列があり，到着客が最短の待ち行列に行くモデルの待ち行列長の定常分布の減少率は容易に推測できるが，$n \geq 3$の場合には証明されていない．同様なことがサービス時間や到着間隔を一般化したジャクソンネットワークについてもいえる．これは，定常分布の存在条件についてもいえることであるが，この安定性条件より，分布の漸近特性の方が問題としてはやさしい場合もあると予想している．

4.4 まとめ

これからの理論研究には，確率以外の他分野の知識がこれまで以上に必要となる．分布の特性を詳しく調べるためには複素関数論，確率過程の極限のためには関数解析，最適化のためには凸解析と変分法などが必要である．これらの基礎理論は完成度が高く優れた教科書がある．研究方法についても触発されることが多い．幅広くこれらを学んで困難な問題に立ち向かってほしい．

5. $GI/G/\infty$ はサービス時間分布に感応する

町原　文明

到着過程がポアソンで無限のサーバ数をもつ待ち行列システム$M/G/\infty$における系内客数の定常分布が，サービス時間の鈍感性（Insensitivity）という特異な性質をもつことはつとに知られている．つまり，その定常分布は$M/M/\infty$のそれに等しく，ポアソン分布となる．鈍感性が成立するには，系内客数過程が可逆となることが必要である．元の過程（順過程）における客の到着は時間軸を逆にたどった逆過程においては退去となるが，可逆性は到着過程と退去過程を確率的に等価にする．この等価性により鈍感性が導かれるわけである．可逆性の成立には到着のポアソン性が必要であり，$GI/G/\infty$においては到着がポアソンでない限り，その鈍感性が雲散霧消する．どこへいってしまうのかを確率的順序付けにより明らかにしたい．

2つの確率変数間の確率的順序付けのためにはいくつかの概念があるが，特にそれらの中で，より変動が大きい（more variable）というものを考える．確率変数Xが確率変数Yより変動が大きい（$X \geq_v Y$と書く）とは，すべての増加凸関数$h(x)$に対して$E[h(X)] \geq E[h(Y)]$が成立することである．それぞれの平均が等しい超指数確率変数H，指数確率変数M，ガンマ確率変数Γ，一定時間Dを確率比較すると$H \geq_v M \geq_v \Gamma \geq_v D$となる．

通信ネットワークの性能評価の分野では，ポアソン到着より変動が大きい故にシステムの性能に深刻な影響を与える到着をバースト到着というが，ここではそれを団塊到着といいかえる．団塊到着とは逆に，ポアソン到着より変動の小さい到着を平滑（smooth）到着とよぶ．以下，到着過程がポアソン到着から団塊到着，あるいは平滑到着に変わるとサービス分布がどのように影響を与え始めるかについて，$GI/G/\infty$の世界の中で考察する．初めに，団塊到着（超指数間隔到着）をもつ$H/G/\infty$に対してサービス時間の変動を小さいほうからD，Γ，M，Hの順に大きくしたときに系内客数の変動がどうなるかをみてみよう．まず，$H/D/\infty$を考える．客の到着時点を一定のサービス時

間ずらすと退去時点になり，退去過程は到着過程に等しくなる．到着の団塊特性は退去にそのまま遺伝する．団塊到着により系内客数が急激に増加するが，一定時間たつと団塊退去が起り，今度は逆に系内客数が急激に減少し，その結果，系内客数が大きく変動することになる．一方，サービス時間が確率的に変動する場合には退去の団塊度が和らぐ．退去過程は平滑化され，到着過程とは異なり緩やかなものとなり，系内客数の変動は小さくなる．この平滑効果は D サービス，Γ サービス，M サービス，H サービスとより変動が大きくなるにつれて顕著なものとなる．その結果，$GI/G/\infty$ の定常状態における系内客数の確率変数を $X_{GI,G}$ と書くと，$X_{H,D} \geq_v X_{H,\Gamma} \geq_v X_{H,M} \geq_v X_{H,H} \geq_v X_{M,G}$ という確率不等式が成立することになる．ここで，サービス時間が重い裾野（heavy tail）をもつ場合には，$X_{H,H}$ と $X_{M,G}$ が確率的に等しくなるという予想をしておく．以上の考察より，$H/G/\infty$ の中で $H/D/\infty$ が最大の系内客数の変動をもち，システム性能上最悪なものとなることがわかる．ここで一つの教訓（冗談）．気分が躁状態にあるときには次々と課題を与える一方で，鬱中にはそれを全く出さない気紛れな教員にどう対処すればいいか？　次々と与えられる課題にいちいち真面目に対応して締切までに完璧に仕上げる努力をすると，脳内システムは $H/D/\infty$ の様相を呈して脳内を並列して走る課題の数が大きく変動する．脳内システムの性能低下は免れ得ない．気紛れには気紛れ，処理（サービス）時間の変動を大きくさせて脳内を H/H (heavy tail)$/\infty$ にもっていけばいいだろう．

それでは逆に平滑到着をもつ $\Gamma/G/\infty$ の場合，鈍感性はどこへ行くのであろうか．団塊到着の場合とは逆に $X_{\Gamma,D} \leq_v X_{\Gamma,\Gamma} \leq_v X_{\Gamma,M} \leq_v X_{\Gamma,H} \leq_v X_{M,G}$ という不等式が成立するはずである．きちんと定期的に課題を出す教員に対してはきちんと対応するのがよい．

システムの良し悪しを判断するには，システムごとの性能を定量的に評価するというより，システム間の確率的な比較のほうが重要視されるようになってきているような気がする（もちろん，システム設備数の算出においては定量的な評価が必須であり，新しい性能評価式を見出すことの意味が薄れたわけではないけれども）．確率比較を使ってシステムを定性的に比較することができれば，本来 OR 研究者がたつべきコンサルタントとしての立場が高まるに違いない．$GI/G/\infty$ の性能比較についてざっと述べてきたが，筆者のいい加減な直感による部分が多い．若い読者の皆さんが，この直感の正否に興味をもたれるなら幸いである．待ちのあるシステム $GI/G/S$ に対しては，述べてきたことが全く成り立たないことを付記する．この問題に対する若い皆様のチャレンジを期待する．

6.　参照局所性の数理的基礎付け

紀　一誠

6.1　問題の設定

離散的な状態空間 $S=\{a_0, a_1, \cdots\}$ あるいは $S_M=\{a_0, a_1, \cdots, a_M\}$ 上の値をとる確率変数列を $R=\{X_n\}_{n=0}^{\infty}$ とし，X_n が値 a_k をとることを，時刻 n で a_k が参照された，ということにする．この R を，参照の局所性（locality）といわれる次の二つの性質を満たすような確率過程として構成するにはどのようにすればよいであろうか．

(a) 時間的局所性：ある時刻である状態を参照した場合には，近い将来その状態をもう一度参照する可能性が高い．

(b) 空間的局所性：ある時刻である状態を参照した場合には，近い将来その状態の近辺の状態を参照する可能性が高い．

6.2　問題の背景

キャッシュ（cache）をはじめとする記憶領域の階層化技術の技術的根拠は，プログラムの動作特性（アドレス参照，ページ参照）には上記(a)，(b)の局所性があるという経験則に基づいている．この経験則は様々な形に表現されているが，最終的には対数線形性をもつ $y=ax^b$ という形に帰着される．ここで例えば，y はキャッシュのミス発生間隔，x はキャッシュサイズであり，a，b は適当な定数である．今までに，この経験則を説明する研究，これを利用して効率的なキャッシュアルゴリズムを提案する研究をはじめとし，おびただしい数の論文が発表されてきた．また，この形はスケール変換不変性をもつことからフラクタルと関連して論じているものもある．しかし，長い歴史と膨大な論文数にもかかわらず，どのような参照構造がこのような局所性を生み出すのか，また時間的局所性と空間的局所性の関係はどのようなものなのか，という本質的な問題に迫る研究は現れていない．

6.3　問題への取り組み

時間的局所性：R として最も簡単な構造である独立参照列（i. i. d. r. v's），次いで簡単な構造であるマルコフ連鎖を仮定した場合について，代表的な指標である以下についての研究がいくつかなされている．

ワーキングセット：時刻 n におけるウインドウサイズ τ のワーキングセットとは，R から生成される確率過程で，時刻 n 以前の τ 時間に参照された状態集合すなわち

$$W(n, \tau)=\{X_k | k \in \max\{0, n-\tau+1\}\} \tag{1}$$

と定義される．また(1)の重複した状態を除いて数えた要素数をワーキングセットサイズといい，$w(n, \tau)=|W(n, \tau)|$ と表す．

LRU (Least-Recently-Used) スタック：状態空間 S_M をもつ R から生成される確率過程 $L(n)=\{L_i(n)\}_{i=1}^{M}$ を時刻 n の LRU スタックという．ただし，$L_i(n)$ は n から時間を遡って i 番目に新しく参照された状態を表す．すなわち $L_1(n)$ は最も最近参照された状態，$L_M(n)$ は最も遠い過去に参照された状態となる．$w(n, \tau)$ や $L_M(n)$ は時間的局所性の評価指標であり，空間的局所性については語らない．

空間的局所性：研究数は少ないが，ランダムウォークモデル $X_{n+1}=X_n+U_n$ を用いた研究がある．ここで，$\{U_n\}_{n=0}^{\infty}$ は空間的局所性を表現する i. i. d. r. v's とし，$E(|U_n|)<\infty$，$V(|U_n|)=\infty$ である．

6.4 問題へのコメント

参照局所性(a)，(b)がプログラムのもつループ構造に起因するであろうことは誰しもが容易に推測することができよう．また，独立参照列や周期性の扱いが苦手なマルコフ連鎖ではその構造をとらえきることは困難であることも容易に想像がつく．研究の中間成果として独立参照列やマルコフ連鎖を導入した結果は貴重ではあるが，本質をとらえたとは言いがたい．二つの性質の異なる参照局所性を統一的にとらえるためには新しい確率過程の構築が必要であり，それは容易な道ではない．もって次世代の研究課題とするゆえんである．