

待ち行列シンポジウム
「確率モデルとその応用」
報文集

PROCEEDINGS
OF
THE QUEUEING SYMPOSIUM
STOCHASTIC MODELS AND THEIR APPLICATIONS

2019 年 1 月 23 日（水）～ 25 日（金）
早稲田大学本キャンパス小野記念講堂（東京都新宿区）

実行委員長：笠原 正治（奈良先端科学技術大学院大学）

主催 待ち行列シンポジウム実行委員会

2018 年度待ち行列シンポジウム「確率モデルとその応用」

2019 年 1 月 23 日 (水) ~25 日 (金)

早稲田大学 本キャンパス 小野記念講堂 (東京都新宿区)

質疑応答 5 分を含め一般 1 人あたり 30 分, 学生 1 人あたり 25 分, ショート 1 人あたり 20 分

* = 発表者 (複数著者の場合)

1 月 23 日 (水)

13:30 開会にあたって 笠原 正治 (奈良先端科学技術大学院大学)

Session W1 (13:40-14:30) 学生セッション 座長: 木村 達明 (大阪大学)

13:40 客のグルーピングのある待ち行列の安定条件: ドリフト条件による導出

* 松尾 容典, 鶴見 星花, 塩田 茂雄 (千葉大学)

14:05 A downlink heterogeneous cellular network model with intra- and inter-tier correlated base station locations

* 小林 拓矢, 三好 直人 (東京工業大学), Dhillon Harpreet S. (Virginia Tech)

Session W2 (14:55-16:25) 一般セッション 座長: 増山 博之 (京都大学)

14:55 A Markov Model of Eusociality at its Origin

* 豊泉 洋 (早稲田大学), Jeremy Field (University of Exeter)

15:25 Correlations of interference in non-Poisson networks under spatially correlated shadowing

木村 達明 (大阪大学)

15:55 SMT ソルバを用いた確率ベトリネットのパーフェクトサンプリングアルゴリズム

* 岡村 寛之, 森原 和也, 土肥 正 (広島大学)

Session W3 (16:50-17:30) ショートペーパーセッション 座長: 加藤 憲一 (神奈川大学)

16:50 ハウスドルフ一次モーメント型トラフィックモデルの性能評価に対する不完全性
高田 寛之 (長崎大学)

17:10 Queues driven by level dependent Levy processes

Czarna Irmina (Wroclaw University of Science and Technology), Perez Jose Luis (CIMAT), Rolski Tomasz (University of Wroclaw), * Yamazaki Kazutoshi (Kansai University)

懇親会 (18:00-20:00) 司会： 増山 博之 (京都大学)

1月24日 (木)

Session T1 (9:30-10:30) ショートペーパーセッション 座長： 岸 康人 (松蔭大学)

- 9:30 真社会性昆虫の基本再生産数
* 石井 諒太郎, 豊泉 洋 (早稲田大学)
- 9:50 乗車時間を考慮したタクシー待ち行列モデルの解析
* 巻野 侑大, Phung-Duc Tuan (筑波大学)
- 10:10 閉園時刻直前のアトラクション巡回法
* 鶴田 裕太郎, 佐藤 真史, 豊泉 洋 (早稲田大学)

Session T2 (10:55-12:35) 学生セッション 座長： 岡村 寛之 (広島大学)

- 10:55 木構造を持つ一製品多工程拡張かんぱんシステムの解析
島先 孝太郎 (東京工業大学)
- 11:20 M/G/1 型マルコフ連鎖に対するレベル増分切断近似の収束性
* 大内 克久, 増山 博之 (京都大学)
- 11:45 待ち行列モデルを用いた宅配ボックスサービスの性能解析
* 日出山 慎人, Phung-Duc Tuan, 岡田 幸彦 (筑波大学)
- 12:10 マルコフ連鎖における条件付き定常分布の線形不等式系による特徴づけ
* 木村 雅俊, 滝根 哲哉 (大阪大学)

12:35-14:05 昼休み

Session T3 (14:05-15:20) 学生セッション 座長： 井上 文彰 (大阪大学)

- 14:05 大数の法則を用いたトールス上の非線形データ拡散モデル
* 渡辺 樹, 豊泉 洋 (早稲田大学)
- 14:30 再試行型待ち行列を用いた単一チャネルコグニティブ無線の解析
* 阿久津 康平, Phung-Duc Tuan (筑波大学)
- 14:55 移動体通信における time-based なハンドオーバー制限のモデルとその解析
* 徳山 喜一, 三好 直人 (東京工業大学)

特別講演 (15:45-16:45) 座長： 笠原 正治 (奈良先端科学技術大学院大学)

待ち行列における到着間隔あるいはサービス時間の変動増加はシステム混雑を増加させるか？

町原 文明 (東京電機大学)

1月25日(金)

Session F1 (9:30-10:30) ショートペーパーセッション 座長：井家 敦(神奈川工科大学)

9:30 客のグルーピングのある待ち行列の安定条件：Foster の定理による導出

* 鶴見 星花, 松尾 容典, 塩田 茂雄(千葉大学)

9:50 短期貸出客を優先処理する図書館の貸し出し図書の待ち行列

* 白木 詩乃, 塩田 茂雄(千葉大学)

10:10 飲食店における客席案内戦略に関する考察

* 野田 脩平, 塩田 茂雄(千葉大学)

Session F2 (10:55-11:55) 一般セッション 座長：小林 正弘(東海大学)

10:55 M/G/1 待ち行列における待ち時間制約と滞在時間制約の比較

井上 文彰(大阪大学)

11:25 メッセージ分割が発生する時のペイロード長が平均応答時間に及ぼす影響—バー
スト的にパケットが到着する場合—

池川 隆司(早稲田大学/東京大学)

11:55 待ち行列研究部会 研究奨励賞 受賞者発表 笠原 正治(奈良先端科学技術大学院
大学)

閉会にあたって 笠原 正治(奈良先端科学技術大学院大学)

目 次

客のグルーピングのある待ち行列の安定条件：ドリフト条件による導出	1
松尾 容典, 鶴見 星花, 塩田 茂雄 (千葉大学)	
A downlink heterogeneous cellular network model with intra- and inter-tier correlated base station locations.....	11
小林 拓矢, 三好 直人 (東京工業大学), Dhillon Harpreet S. (Virginia Tech)	
A Markov Model of Eusociality at its Origin	13
豊泉 洋 (早稲田大学), Jeremy Field (University of Exeter)	
Correlations of interference in non-Poisson networks under spatially correlated shadowing	23
木村 達明 (大阪大学)	
SMT ソルバを用いた確率ペトリネットのパーフェクトサンプリングアルゴリズム	33
岡村 寛之, 森原 和也, 土肥 正 (広島大学)	
ハウスドルフ一次モーメント型トラフィックモデルの性能評価に対する不完全性	42
高田 寛之 (長崎大学)	
Queues driven by level dependent Levy processes.....	44
Czarna Irmina (Wroclaw University of Science and Technology), Perez Jose Luis (CIMAT), Rolski Tomasz (University of Wroclaw), Yamazaki Kazutoshi (Kansai University)	
真社会性昆虫の基本再生産数	46
石井 諒太郎, 豊泉 洋 (早稲田大学)	
乗車時間を考慮したタクシー待ち行列モデルの解析	48
巻野 侑大, Phung-Duc Tuan (筑波大学)	
閉園時刻直前のアトラクションの選び方	50
鶴田 裕太郎, 佐藤 真史, 豊泉 洋 (早稲田大学)	
木構造を持つ一製品多工程拡張かんぱんシステムの解析	52
島先 孝太郎 (東京工業大学)	
M/G/1 型マルコフ連鎖に対するレベル増分切断近似の漸近公式	62
大内 克久, 増山 博之 (京都大学)	
待ち行列モデルを用いた宅配ボックスサービスの性能解析	71
日出山 慎人, Phung-Duc Tuan, 岡田 幸彦 (筑波大学)	

マルコフ連鎖における条件付き定常分布の線形不等式系による特徴づけ	81
木村 雅俊, 滝根 哲哉 (大阪大学)	
大数の法則を用いたトーラス上のデータ拡散モデル	91
渡辺 樹, 豊泉 洋 (早稲田大学)	
再試行型待ち行列を用いた単一チャネルコグニティブ無線の解析	99
阿久津 康平, Phung-Duc Tuan (筑波大学)	
移動体通信における Time-based なハンドオフ制限のモデルとその解析	109
徳山 喜一, 三好 直人 (東京工業大学)	
客のグルーピングのある待ち行列の安定条件: Foster の定理による導出	119
鶴見 星花, 松尾 容典, 塩田 茂雄 (千葉大学)	
短期貸出客を優先処理する図書館の貸出図書の待ち行列	121
白木 詩乃, 塩田 茂雄 (千葉大学)	
飲食店における客席案内戦略に関する考察	123
野田 脩平, 塩田 茂雄 (千葉大学)	
M/G/1 待ち行列における待ち時間制約と滞在時間制約の比較	125
井上 文彰 (大阪大学)	
メッセージ分割が発生する時のペイロード長が平均応答時間に及ぼす影響—バースト的にパケットが到着する場合—	133
池川 隆司 (早稲田大学/東京大学)	

客のグルーピングのある待ち行列の安定条件：ドリフト条件による導出

松尾 容典* 鶴見 星花† 塩田 茂雄*

* 千葉大学 大学院融合理工学府

† 千葉大学 工学部

概要：本稿では、1つのサービス窓口に対して、複数人がグループとなって同時にサービスを受けることが可能な（グループになることを拒否する客も考慮する）モデルを考える．このモデルをグルーピングモデルと呼ぶ．グルーピングモデルは、ブロック行列がレベル依存であるため、安定条件を簡単に求めることはできないが、サービス規律を変えた2種類のモデルと利用率を比較することで、安定条件の導出を行った．結果的に、グルーピングモデルの安定条件は $\rho_1 + \rho_2/N < 1$ で表すことができることが分かった．

1. まえがき

待ち行列理論では、1つのサービス窓口に対して、1つの客（団体）がサービスを受けるモデルを対象として扱われてきた．しかし、実際のサービスにおいては、シェアリングが許されるケースも見受けられる．例として、ジェットコースターの車両1台に対して、複数の客が同時にサービスを受ける例や、飲食店において、2人用のテーブルに対して別々の客が相席をするような例が挙げられる．このうち、後者のようにシェアリングの可否が客により異なるモデルについての研究はなされていない．よって、本稿においては、1つのサービス窓口に対して、複数人がグループとなって同時にサービスを受けることが可能な待ち行列モデル（以降グルーピングモデル）を構築し、シェアリングの可否が客により異なるモデルについて考察する．ただし、グルーピングモデルはドリフト条件を用いた安定状態の導出が困難であるため、サービス規律を変えた2種類のモデルを用意し、それらの安定条件をドリフト条件を用いて導出することで、その安定条件を元にグルーピングモデルの安定条件の導出を行っていく．

2. 条件設定

本稿では、M/M/1モデルを元に、2種類のタイプの客が到着する待ち行列モデルを考える．2種類のタイプの客をそれぞれタイプ1、タイプ2と呼び、それぞれ到着率 λ_1 および λ_2 のポアソン過程に従って到着する．タイプ1の客はシェアリングを拒否する客であり、1人でサービスを受けることになる．一方、タイプ2の客はシェアリングを許可する客であり、最大 N 人のグループを組んで、サービスを受けることになる．タイプ1の客は、到着時に行列の最後尾に並ぶ．タイプ2の客は、行列内に N 人未満のグループ（以降、未完成グループ）が存在すれば、そのグループに加入し、未完成グループが存在しなければ、行列の最後尾に1人で並ぶ．客のサービスは行列の先頭にいる客（グループ）から行われ、そのサービス時間は客のタイプやグループの構成人数に関わらず平均 $1/\mu$ の指数分布に従うとする．以降、 $\rho_1 := \lambda_1/\mu, \rho_2 := \lambda_2/\mu$ とする．

注 2.1. タイプ2の未完成グループがサービスを受けている最中にタイプ2の客の到着した場合、到着客のサービス中の未完成グループへの加入を許す場合と許さない場合の2通りのサービス規律が存在する．本稿では、加入を許さない場合を想定して議論を進めるが、加入を許す場合でも本稿の結論は変わらない．

3. グルーピングモデル

時刻 t における、系内客数を $L(t)$ で表す．ただし、グループは、何人でグループを組んでいようと、1人としてカウントする．また、未完成グループの待ち行列の先頭からの位置を $I(t)$ 、未完成グループの構成人数を $J(t)$ で表す．ただし、未完成グループが待ち行列に存在しない場合は $I(t) = J(t) = 0$ とする．また未完成グループに対してサービスが始まった場合、（構成人数が N 人未満であっても）サービス開始以降そのグループは完成グループとして扱うものとする．これらの確率変数を用いた確率過程 $\{L(t), I(t), J(t); t \geq 0\}$ は連続時間マルコフ連鎖となる．待ち行列に並んでいる人数が $l (l \geq 0)$ 、未完成グループの位置が $i (0 \leq i \leq l)$ 、未完成グループの構成人数が $j (0 \leq j \leq N-1)$ 、サービスを受けている客がいる状態を、 (l, i, j) で表し、状態 (l, i, j) の定常状態分布を $\pi_{l,i,j}$ で表す．さらに、待ち行列に並んでいる人数が $l (> 1)$ の定常分布ベクトルを辞書式に $\pi_l = (\pi_{l,0,0}, \pi_{l,2,1}, \dots, \pi_{l,2,N-1}, \dots, \pi_{l,l,1}, \dots, \pi_{l,l,N-1})$ と定める¹．なお、 $\pi_0 = \pi_{0,0,0}, \pi_1 = \pi_{1,0,0}$ である．定

¹ 本モデルにおいては、未完成グループは先頭の位置には存在しないので、 $I(t) \geq 2$ である．

常分布の大域平衡方程式は次のように行列形式に書き下すことができる。

$$\begin{aligned}\pi_0 D_0^{(0)} + \pi_1 D_{-1}^{(0)} &= 0, \\ \pi_{l-1} D_{+1}^{(l)} + \pi_l D_0^{(l)} + \pi_{l+1} D_{-1}^{(l)} &= \mathbf{0}. \quad (l = 1, 2, \dots)\end{aligned}$$

大域平衡方程式に表れるブロック行列は次の通りである。

$$\begin{aligned}D_{+1}^{(l)} &= \begin{pmatrix} \lambda_1 & \mathbf{0} & \cdots & \mathbf{0} & \lambda_2 \mathbf{e}_1 \\ \mathbf{0}^\top & \lambda_1 I_{N-1} & O & \cdots & O \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}^\top & \cdots & O & \lambda_1 I_{N-1} & O \end{pmatrix}, \quad D_{-1}^{(l)} = \begin{pmatrix} \mu & \mathbf{0} & \cdots & \mathbf{0} \\ \mu \mathbf{1}^\top & O & \cdots & O \\ \mathbf{0}^\top & \mu I_{N-1} & \ddots & O \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}^\top & \cdots & O & \mu I_{N-1} \end{pmatrix}, \\ D_0^{(l)} &= -(\lambda_1 + \lambda_2 + \mu) I_{l(N-1)+1} + \begin{pmatrix} 0 & O & \cdots & O \\ \lambda_2 \mathbf{e}_{N-1}^\top & \lambda_2 I_{N-1}^+ & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_2 \mathbf{e}_{N-1}^\top & O & \cdots & \lambda_2 I_{N-1}^+ \end{pmatrix}, \\ D_0^{(0)} &= -(\lambda_1 + \lambda_2), \quad D_{-1}^{(0)} = \mu, \quad D_{+1}^{(1)} = \lambda_1 + \lambda_2, \quad D_0^{(1)} = -(\lambda_1 + \lambda_2 + \mu), \quad I_n^+ = \begin{pmatrix} \mathbf{0}^\top & I_{n-1} \\ 0 & \mathbf{0} \end{pmatrix}.\end{aligned}$$

ここで、 $\mathbf{0}$ は全ての要素が 0 の横ベクトル、 $\mathbf{1}$ は全ての要素が 1 の横ベクトル、 $\mathbf{e}_n$ は $\mathbf{0}$ の n 番目の要素のみを 1 に置き換えた横ベクトル、 I_n は n 次の単位行列を表す。なお、

$$\begin{cases} D_{-1}^{(l)} : (l(N-1)+1) \times ((l-1)(N-1)+1) \text{ の行列} & (l = 1, 2, \dots) \\ D_0^{(l)} : ((l-1)(N-1)+1) \text{ 次の正方行列} & (l = 2, 3, \dots) \\ D_{+1}^{(l)} : ((l-2)(N-1)+1) \times ((l-1)(N-1)+1) \text{ の行列} & (l = 2, 3, \dots) \end{cases}$$

である。さらに $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots)$ とし、 Q を次で定義することで、大域平衡方程式は $\boldsymbol{\pi} Q = \mathbf{0}$ と書ける。

$$Q = \begin{pmatrix} D_0^{(0)} & D_{+1}^{(1)} & O & \cdots & \cdots \\ D_{-1}^{(0)} & D_0^{(1)} & D_{+1}^{(2)} & \ddots & \\ O & D_{-1}^{(1)} & D_0^{(2)} & D_{+1}^{(l)} & \ddots \\ \vdots & \ddots & D_{-1}^{(2)} & D_0^{(l)} & \ddots \\ \vdots & & \ddots & \ddots & \ddots \end{pmatrix}.$$

このように、グルーピングモデルは準出生死滅過程であるが、ブロック行列のサイズがレベル依存であるため、このままではドリフト条件を用いて安定条件を導出するには不便な形となっている。そこで、本稿では、到着した客のサービス規律を変えた 2 つのモデルである、タイプ 2 優先 (K -priority) モデルとタイプ 2 非優先 (Non-priority) モデルと利用率を比較することで、グルーピングモデルの安定条件の検討を行う。

4. タイプ 2 優先 (K -priority) モデル

最初に定義したグルーピングモデルは、ブロック行列がレベル依存であることが、安定条件の導出を難しくしていた。そこで、サービスを受けていない未完成グループが存在しないときに到着したタイプ 2 の客を、待ち行列の先頭から K 番目に並ばせるサービス規律を考える。ただし、待ち行列に並んでいる人数が K 未満の場合、従来通り待ち行列の最後尾に並ばせる。本稿では、このようなサービス規律を持つモデルをタイプ 2 優先モデルと呼ぶ。以降、略称として K -priority モデル、もしくは KP モデルと呼ぶ。 K -priority モデルにおける定常分布の大域平衡方程式は、次のように行列形式に書き下すことができる。

$$\begin{aligned}\pi_0 D_0^{(0)} + \pi_1 D_{-1}^{(0)} &= 0, \\ \pi_{l-1} D_{+1}^{(l)} + \pi_l D_0^{(l)} + \pi_{l+1} D_{-1}^{(l)} &= \mathbf{0}, \quad (l = 1, \dots, K-1) \\ \pi_{K-1} D_{+1}^{(K)} + \pi_K D_0^{(K)} + \pi_{K+1} A_{-1}^{(K)} &= \mathbf{0}, \\ \pi_{l-1} A_{+1}^{(K)} + \pi_l D_0^{(K)} + \pi_{l+1} A_{-1}^{(K)} &= \mathbf{0}. \quad (l = K+1, K+2, \dots)\end{aligned}$$

ここで $D_{+1}^{(l)}$, $D_0^{(l)}$, $D_{-1}^{(l)}$ は、いずれもグルーピングモデルで定義した行列である。また、 $A_{+1}^{(K)}$ と $A_{-1}^{(K)}$ は $((K-1)(N-1)+1) \times ((K-1)(N-1)+1)$ の以下の正方形行列である。

$$A_{+1}^{(K)} = \begin{pmatrix} \lambda_1 & \mathbf{0} & \cdots & \mathbf{0} & \lambda_2 e_1 \\ \mathbf{0}^\top & \lambda_1 I_{N-1} & O & \cdots & O \\ \mathbf{0}^\top & O & \lambda_1 I_{N-1} & \cdots & O \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}^\top & O & \cdots & O & \lambda_1 I_{N-1} \end{pmatrix}, \quad A_{-1}^{(K)} = \begin{pmatrix} \mu & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mu \mathbf{1}^\top & O & \cdots & \cdots & O \\ \mathbf{0}^\top & \mu I_{N-1} & \ddots & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}^\top & \cdots & O & \mu I_{N-1} & O \end{pmatrix}.$$

なお、 K -priority モデルでは、系内客数が l 人である定常分布ベクトル π は、 $l \leq K$ のときはグルーピングモデルのときと同じであるが、 $l > K$ のとき、以下のように長さ $(K-1)(N-1)+1$ の横ベクトルとなる。

$$\pi_l = (\pi_{l,0,0}, \pi_{l,2,1}, \dots, \pi_{l,2,N-1}, \dots, \pi_{l,K,1}, \dots, \pi_{l,K,N-1}) \quad (l = K, K+1, \dots)$$

K -priority モデルは、 $A_{+1}^{(K)}$, $D_0^{(K)}$, $A_{-1}^{(K)}$ がいずれも $(K-1)(N-1)+1$ 次の正方形行列となることから、確率過程 $\{L(t), I(t), J(t); t \geq 0\}$ は、系内客数が K 人以上のとき、ブロック行列の大きさが $((K-1)(N-1)+1) \times ((K-1)(N-1)+1)$ で固定する準出生死滅過程となり、その定常分布は次で与えられる。

$$\pi_l = \pi_K R^{l-K}, \quad (l = K+1, K+2, \dots)$$

ただし、 R は公比行列であり、次の方程式を満たす非負最小解として与えられる。

$$A_{+1}^{(K)} + R D_0^{(K)} + R^2 A_{-1}^{(K)} = O.$$

また $\pi_0, \dots, \pi_K$ は次の解である。

$$(\pi_0, \dots, \pi_K) \begin{pmatrix} D_0^{(0)} & D_{+1}^{(1)} & O \\ D_{-1}^{(0)} & D_0^{(1)} & D_{+1}^{(2)} \\ & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & \ddots \\ & & & D_{-1}^{(K-2)} & D_0^{(K-1)} & D_{+1}^{(K)} \\ & & & O & D_{-1}^{(K-1)} & D_0^{(K)} + R A_{-1}^{(K)} \end{pmatrix} = (\mathbf{0}, \dots, \mathbf{0}),$$

$$\pi_0 \mathbf{1}^\top + \dots + \pi_K (I - R)^{-1} \mathbf{1}^\top = 1.$$

K -priority モデルの安定条件は、

$$D^{(K)} \stackrel{\text{def}}{=} A_{+1}^{(K)} + D_0^{(K)} + A_{-1}^{(K)} = \begin{pmatrix} -\lambda_2 & \mathbf{0} & \cdots & \mathbf{0} & \lambda_2 e_1 \\ \lambda_2 e_{N-1}^\top + \mu \mathbf{1}^\top & C & O & \cdots & O \\ \lambda_2 e_{N-1}^\top & \mu I_{N-1} & C & \ddots & O \\ \vdots & \ddots & \ddots & \ddots & O \\ \lambda_2 e_{N-1}^\top & O & \cdots & \mu I_{N-1} & C \end{pmatrix},$$

$$C \stackrel{\text{def}}{=} \lambda_2 (I_{N-1}^+ - I_{N-1}) - \mu I_{N-1}.$$

の定常分布 $\tilde{\pi} = (\tilde{\pi}_0, \tilde{\pi}_{1,1}, \dots, \tilde{\pi}_{K-1,N-1})$ が次のドリフト条件 [1] を満たすかどうかで定まる。

$$\tilde{\pi} A_{+1}^{(K)} \mathbf{1}^\top < \tilde{\pi} A_{-1}^{(K)} \mathbf{1}^\top.$$

ドリフト条件は $\rho_1 + \tilde{\pi}_0 \rho_2 < 1$ と等価である。

補題 4.1. $\tilde{\pi}_0, \tilde{\pi}_{i,j}$ は以下で定まる。次の補題が成立する（証明は付録に示す）。

$$\tilde{\pi}_{i,j} = \binom{K-i+j-2}{j-1} (1-q)^{K-1-i} q^j \tilde{\pi}_0,$$

$$\tilde{\pi}_0 = \left(N - (1+\rho_2)^{-(K-1)} \sum_{j=1}^{N-1} \sum_{k=1}^j \left(\frac{\rho_2}{1+\rho_2} \right)^{k-1} \binom{K+k-3}{k-1} \right)^{-1}.$$

補題 4.1 より, K -priority モデルの安定条件が以下のように得られる.
 定理 4.2. 次の式を満たすとき, K -priority モデルは安定である.

$$\rho_1 + \frac{\rho_2}{N - (1 + \rho_2)^{-(K-1)} \sum_{j=1}^{N-1} \sum_{k=1}^j \left(\frac{\rho_2}{1 + \rho_2} \right)^{k-1} \binom{K+k-3}{k-1}} < 1.$$

5. タイプ 2 非優先 (Non-priority) モデル

K -priority モデルでは, サービスを受けていない未完成グループが存在しないときに到着したタイプ 2 の客のを, 待ち行列の先頭から K 番目に並ばせるサービス規律を考えていた. このサービス規律は K が小さければ小さいほど, タイプ 2 の客がペアを組むことができる可能性が低いという特徴を持っている. そこで, 逆に最大限タイプ 2 の客がペアを組むことができる可能性が高いサービス規律として, 未完成グループを『常に』待ち行列の最後尾に並ばせる (サービスを受け始めた未完成グループはこの限りではない) サービス規律を考える. 本稿では, このようなサービス規律を持つモデルをタイプ 2 非優先モデルと呼ぶ. 以降, 略称として Non-priority モデル, np モデルと呼ぶ. Non-priority モデルにおける定常分布の大域平衡方程式は, 次のように行列形式に書き下すことができる.

$$\begin{aligned} \pi_0 D_0^{(0)} + \pi_1 D_{-1}^{(0)} &= 0, \\ \pi_0 D_{+1}^{(1)} + \pi_1 D_0^{(1)} + \pi_2 D_{-1}^{(1)} &= 0, \\ \pi_1 D_{+1}^{(2)} + \pi_2 D_0^{(2)} + \pi_3 B_{-1}^{(2)} &= 0, \\ \pi_{l-1} A_{+1}^{(2)} + \pi_l D_0^{(2)} + \pi_{l+1} B_{-1}^{(2)} &= 0. \quad (l = 3, 4, \dots) \end{aligned}$$

ここで $B_{-1}^{(2)} \stackrel{\text{def}}{=} \mu I_N$ である. Non-priority モデルでは, 系内客数 l が 2 以上のときの定常分布ベクトル π_l は以下のように長さ N の横ベクトルとなる.

$$\pi_l = (\pi_{l,0,0}, \pi_{l,2,1}, \dots, \pi_{l,2,N-1}) \quad (l = 2, 3, \dots)$$

$A_{+1}^{(2)}, D_0^{(2)}, B_{-1}^{(2)}$ はいずれも N 次の正方行列であることから, Non-priority モデルは確率過程 $\{L(t), I(t), J(t); t \geq 0\}$ は, 系内客数が 2 以上のとき, ブロック行列の大きさが $N \times N$ で固定される準出生死滅過程となり, その定常分布は次のように与えられる.

$$\pi_l = \pi_2 R^{l-2}, \quad l = 3, 4, \dots$$

R は R は公比行列であり, 次の方程式を満たす非負最小解として与えられる.

$$A_{+1}^{(2)} + R D_0^{(2)} + R^2 B_{-1}^{(2)} = O.$$

また, π_0, π_1, π_2 は次の解である.

$$(\pi_0, \pi_1, \pi_2) \begin{pmatrix} D_0^{(0)} & D_{+1}^{(1)} & \mathbf{0} \\ D_{-1}^{(0)} & D_0^{(1)} & D_{+1}^{(2)} \\ \mathbf{0}^\top & D_{-1}^{(1)} & D_0^{(2)} + R B_{-1}^{(2)} \end{pmatrix} = (0, 0, \mathbf{0}),$$

$$\pi_0 + \pi_1 + \pi_2 (I - R)^{-1} \mathbf{1}^\top = 1.$$

Non-priority 処理のもとでの安定条件は

$$D^{(K)} \stackrel{\text{def}}{=} A_{+1}^{(2)} + D_0^{(2)} + B_{-1}^{(1)} = \begin{pmatrix} -\lambda_2 & \lambda_2 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \ddots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & -\lambda_2 & \lambda_2 \\ \lambda_2 & 0 & \dots & 0 & -\lambda_2 \end{pmatrix},$$

の定常分布 $\mathbf{x} = (1/N, \dots, 1/N)$ が次のドリフト条件 [1] を満たすかどうかで定まる.

$$\mathbf{x} A_{+1}^{(2)} \mathbf{1}^\top < \mathbf{x} B_{-1}^{(2)} \mathbf{1}^\top.$$

これより, 次の結果が得られる.

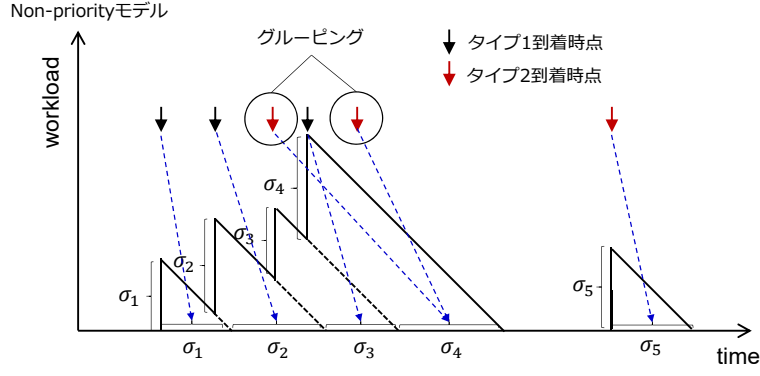


図 1: Workload の変化 (Non-priority モデル)

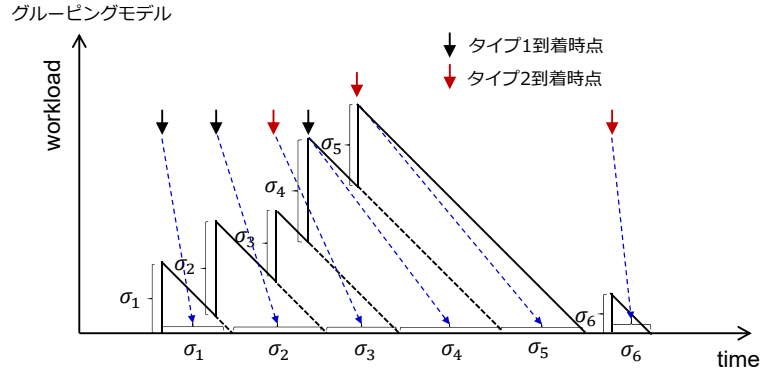


図 2: Workload の変化 (グルーピングモデル)

定理 5.1. 次の式を満たすとき, Non-priority モデルは安定である.

$$\rho_1 + \frac{\rho_2}{N} < 1.$$

6. 利用率の比較, 安定条件の導出

本章では, 当初のポアソン到着, 指数サービス時間の仮定を緩め, 一般到着, 一般サービス時間 (ただし再生性は仮定する: $G_1+G_2/GI/1$ キュー) のもとで, グルーピングモデル, K -priority モデル, Non-priority モデルでの処理をサンプルパス毎に比較しながら, グルーピングモデルの安定条件を導出する. 比較を容易にするため, サンプル空間 Ω 内の任意の要素 $\omega \in \Omega$ に対して, 時刻 0 以降のタイプ 1 およびタイプ 2 の客の到着時刻列 $\{t_n^{(1)}(\omega)\}_{n=1}^{\infty}$, $\{t_n^{(2)}(\omega)\}_{n=1}^{\infty}$, および客のサービス時間列 $\{\sigma_n(\omega)\}_{n=1}^{\infty}$ が定まる確率空間 $(\Omega, \mathcal{F}, P)$ を考える. ここで, σ_n は (時刻 0 以降に) 窓口で n 番目にサービスを受ける客のサービス時間を表す. この確率空間では, 客のサービス時間は客に付随するものではなく, 窓口で何番目にサービスを受けるかで決まる点に留意する. 図 1 に $N = 2$ の場合の Non-priority モデルにおける workload の変化を示す. 図 1 において, 3 番目にタイプ 2 の客が到着すると, その時点で待ち行列システムの workload は σ_3 だけ増加するが, σ_3 は到着したタイプ 2 の客が窓口で受けるサービス時間ではない. 4 番目にタイプ 1 の客が到着すると, Non-priority モデルにおいてはその客の処理が先に行われ, σ_3 は 4 番目に到着したタイプ 1 の客が窓口で受けるサービス時間となる. 3 番目に到着したタイプ 2 の客のサービス時間は σ_4 に等しい. 参考までに, 図 2 にグルーピングモデルにおける workload の変化を示す.

本章では, グルーピングモデル, K -priority モデル, Non-priority モデルの利用率に着目する. 一般に, 利用率 ρ は次の式で求められることが知られている.

$$\rho = E[\sigma] \times \lim_{t \rightarrow \infty} \frac{E[N(t)]}{t} \quad (1)$$

ただし, $E[\sigma]$ は平均サービス時間を, $N(t)$ は時刻 0 から $t \geq 0$ までの累積到着人数を表し, $N(0) = 0$ を仮定する. 平均サービス時間 $E[\sigma]$ はどのモデルにおいても等しいので, $E[N(t)]$ が利用率 ρ の大きさを左右するただ 1 つのパラメータである. ここで, サービス窓口から見た $E[N(t)]$ について考えると, $N(t)$ は, タイプ 1

の累積客数と、到着時にグルーピングができなかったタイプ2の累積客数の和となる。言い換えると、タイプ2の客が到着時にグルーピングできた場合、系内客数 $L(t)$ は増えていないため、到着時にグルーピングできたタイプ2の客は到着時に間引かれていると考えることができる。よって $N_1(t)$ を時刻0から $t \geq 0$ までのタイプ1の累積到着人数、 $\tilde{N}_2(t)$ をグルーピングモデルにおいて時刻0から $t \geq 0$ までの、到着時にグルーピングできなかったタイプ2の客の累積到着人数とおき、 $N_1(0) = \tilde{N}_2(0) = 0$ を仮定することで、次の式が成り立つ。

$$N(t) = N_1(t) + \tilde{N}_2(t) \quad (2)$$

したがって、グルーピングモデルにおける使用率 ρ は

$$\rho = E[\sigma] \times \lim_{t \rightarrow \infty} \frac{E[N_1(t) + \tilde{N}_2(t)]}{t}.$$

となる。同様に、 K -priority モデルおよび Non-priority モデルにおいて、時刻 $t \geq 0$ までにグルーピングされなかったタイプ2の客の累積到着人数を $\tilde{N}_2^{kp}(t)$ および $\tilde{N}_2^{np}(t)$ とおき、 K -priority モデルおよび Non-priority モデルにおける使用率をそれぞれ ρ^{kp} および ρ^{np} とおくと、

$$\rho^{kp} = E[\sigma] \times \lim_{t \rightarrow \infty} \frac{E[N_1(t) + \tilde{N}_2^{kp}(t)]}{t}, \quad \rho^{np} = E[\sigma] \times \lim_{t \rightarrow \infty} \frac{E[N_1(t) + \tilde{N}_2^{np}(t)]}{t},$$

となる。

6.1. グルーピングモデルと K -priority モデルの比較

比較のために、次の補題を使用する（証明は付録で述べる）。

補題 6.1. $W(t)$ および $W^{kp}(t)$ を、グルーピングモデルおよび K -priority モデルにおける時刻 t での workload とする。このとき $t_n^{(2)} \leq t$ を満たす全ての n について

$$N_2^{kp}(t_n^{(2)}) \geq N_2(t_n^{(2)}), \quad (3)$$

が成立し、かつ $N_2^{kp}(t) = N_2(t)$ であるならば、 $W^{kp}(t) \leq W(t)$..

この補題より、次の定理が成り立つことが示される。

定理 6.2. グルーピングモデルと K -priority モデルについて、サンプルパスごとに次の式が成り立つ。

$$\tilde{N}_2(t) \leq \tilde{N}_2^{kp}(t).$$

証明. まず n 番目のタイプ2の客が到着した時刻を表す変数 $t_n^{(2)}$ 、グルーピングモデル、 K -priority モデルにおける時刻 t での workload を表す変数 $W(t), W^{kp}(t)$ 、 K -priority モデルにおける時刻 t での未完成グループの構成人数を表す変数 $J^{kp}(t)$ を導入する。

$\tilde{N}_2(t), \tilde{N}_2^{kp}(t)$ が増加する可能性のあるタイミングは、タイプ2の客が到着したときに限られるので、 $\tilde{N}_2(t) \leq \tilde{N}_2^{kp}(t)$ を証明するためには、 $\tilde{N}_2(t_n^{(2)}) \leq \tilde{N}_2^{kp}(t_n^{(2)})$ が示せば十分であり、これを背理法を用いて示す。 $\tilde{N}_2(t_n^{(2)}) \leq \tilde{N}_2^{kp}(t_n^{(2)})$ でないとすると、初めて $\tilde{N}_2(t_n^{(2)}) > \tilde{N}_2^{kp}(t_n^{(2)})$ となる $n = k$ が存在することになる。言い換えれば次の式を満たす自然数 k が存在する。

$$\tilde{N}_2(t_{k-1}^{(2)}) = \tilde{N}_2^{kp}(t_{k-1}^{(2)}) \quad , \quad \begin{cases} \tilde{N}_2(t_k^{(2)}) &= \tilde{N}_2(t_{k-1}^{(2)}) + 1 \\ \tilde{N}_2^{kp}(t_k^{(2)}) &= \tilde{N}_2^{kp}(t_{k-1}^{(2)}) \end{cases} \quad (4)$$

この式は時刻 $t_k^{(2)}$ において、グルーピングモデルではグルーピングが生じておらず、 K -priority モデルではグルーピングが生じていることを表している。ここで、時刻 $t_k^{(2)}$ において、グルーピングモデルではグルーピングが生じていないことに着目すると、グルーピングモデルについて、次の二つの場合のどちらかが成り立つ。以降、それぞれの場合について考える。

Case1: 時刻 $t_{k-1}^{(2)}$ に到着したタイプ2の客はグルーピングが生じ、そのグループがちょうど N 人になった

Case2: 時刻 $t_{k-1}^{(2)}$ に到着したタイプ2の客は時刻 $t_k^{(2)}$ 以前に未完成グループのまま退去する

まず Case1 について考える. 条件より $J(t_{k-1}^{(2)}) = 0, J(t_{k-2}^{(2)}) = N-1, \dots, J(t_{k-N}^{(2)}) = 1$ となるため, 次の式が成り立つ.

$$\tilde{N}(t_{k-1}^{(2)}) = \tilde{N}(t_{k-1-j}^{(2)}) = \tilde{N}(t_{k-N}^{(2)}) = \tilde{N}(t_{k-N-1}^{(2)}) + 1 \quad (j = 1, \dots, N-1) \quad (5)$$

一方で, K -priority モデルにおいて, 時刻 $t_k^{(2)}$ に到着したタイプ 2 の客はグルーピングできていることから, $J^{kp}(t_{k-1}^{(2)}) = m \leq N-1$ と表すことができる. ただし, m は自然数である. すると, $J^{kp}(t_{k-1}^{(2)}) = m, J^{kp}(t_{k-2}^{(2)}) = m-1, \dots, J^{kp}(t_{k-m}^{(2)}) = 1$ となるため, 次の式が成り立つ.

$$\tilde{N}^{kp}(t_{k-1}^{(2)}) = \tilde{N}^{kp}(t_{k-m}^{(2)}) = \tilde{N}^{kp}(t_{k-m-1}^{(2)}) + 1 \quad (6)$$

よって, (4), (5), (6) より, $\tilde{N}(t_{k-m-1}^{(2)}) > \tilde{N}^{kp}(t_{k-m-1}^{(2)})$ となるが, これは $n = k$ で初めて $\tilde{N}_2(t_n^{(2)}) > \tilde{N}_2^{kp}(t_n^{(2)})$ となることに矛盾する. よって Case1 はあり得ないことが分かる.

次に Case2 について考える. $J(t_{k-1}^{(2)}) = l, J^{kp}(t_{k-1}^{(2)}) = m$ とおく. ただし, l, m は自然数である. すると, $J(t_{k-1}^{(2)}) = l, J(t_{k-2}^{(2)}) = l-1, \dots, J(t_{k-l}^{(2)}) = 1$ かつ $J^{kp}(t_{k-1}^{(2)}) = m, J^{kp}(t_{k-2}^{(2)}) = m-1, \dots, J^{kp}(t_{k-m}^{(2)}) = 1$ となるため, 次の式が成り立つ.

$$\begin{aligned} \tilde{N}(t_{k-1}^{(2)}) &= \tilde{N}(t_{k-1-j}^{(2)}) = \tilde{N}(t_{k-l}^{(2)}) = \tilde{N}(t_{k-l-1}^{(2)}) + 1 \quad (j = 1, \dots, l-1) \\ \tilde{N}^{kp}(t_{k-1}^{(2)}) &= \tilde{N}^{kp}(t_{k-1-j}^{(2)}) = \tilde{N}^{kp}(t_{k-m}^{(2)}) = \tilde{N}^{kp}(t_{k-m-1}^{(2)}) + 1 \quad (j = 1, \dots, m-1) \end{aligned}$$

$l > m$ のとき, $\tilde{N}(t_{k-1}^{(2)}) = \tilde{N}(t_{k-m-1}^{(2)})$, $\tilde{N}^{kp}(t_{k-1}^{(2)}) = \tilde{N}^{kp}(t_{k-m-1}^{(2)}) + 1$ となり, $\tilde{N}(t_{k-m-1}^{(2)}) > \tilde{N}^{kp}(t_{k-m-1}^{(2)})$ となるが, これは $n = k$ で初めて $\tilde{N}_2(t_n^{(2)}) > \tilde{N}_2^{kp}(t_n^{(2)})$ となることに矛盾する. $l \leq m$ のとき, $t_{k-l-1}^{(2)}$ に到着した客に着目すると, グルーピングモデルではこの客は行列の最後尾に加入することから, 補題 6.1 より, $W^{kp}(t_{k-l-1}^{(2)}) \leq W(t_{k-l-1}^{(2)})$ とあわせて考えると, $t_{k-l-1}^{(2)}$ に到着した客は K -priority モデルの方がグルーピングモデルよりも先に退去することになる. しかし, $t_k^{(2)}$ では, グルーピングモデルにおいて $t_{k-l-1}^{(2)}$ に到着した客が存在せず, K -priority モデルにおいて, $t_{k-l-1}^{(2)}$ に到着した客が存在することと矛盾する. よって Case2 はあり得ないことが分かる. $\square$

6.2. グルーピングモデルと Non-priority の比較

比較のために, 次の補題を使用する (証明は付録で述べる).

補題 6.3. $W(t)$ および $W^{np}(t)$ を, グルーピングモデルおよび Non-priority モデルにおける時刻 t での workload とする. このとき $t_n^{(2)} \leq t$ を満たす全ての n について

$$N_2^{np}(t_n^{(2)}) \leq N_2(t_n^{(2)}), \quad (7)$$

が成立し, かつ $N_2^{np}(t) = N_2(t)$ であるならば, $W^{np}(t) \geq W(t)$.

この補題より, 次の定理が成り立つことが示される.

定理 6.4. グルーピングモデルと Non-priority モデルについて, サンプルパスごとに次の式が成り立つ.

$$\tilde{N}_2(t) \geq \tilde{N}_2^{np}(t).$$

証明. n 番目のタイプ 2 の客が到着した時刻を表す変数 $t_n^{(2)}$, グルーピングモデル, Non-priority モデルにおける時刻 t での workload を表す変数 $W(t), W^{np}(t)$, Non-priority モデルにおける時刻 t での未完成グループの構成人数を表す変数 $J^{np}(t)$ を導入する.

$\tilde{N}_2(t), \tilde{N}_2^{np}(t)$ が増加する可能性のあるタイミングは, タイプ 2 の客が到着したときに限られるので, $\tilde{N}_2(t) \geq \tilde{N}_2^{np}(t)$ を証明するためには, $\tilde{N}_2(t_n^{(2)}) \geq \tilde{N}_2^{np}(t_n^{(2)})$ が示せば十分であり, これを背理法を用いて示す. $\tilde{N}_2(t_n^{(2)}) \geq \tilde{N}_2^{np}(t_n^{(2)})$ でないとすると, 初めて $\tilde{N}_2(t_n^{(2)}) < \tilde{N}_2^{np}(t_n^{(2)})$ となる $n = k$ が存在することになる. 言い換えれば次の式を満たす自然数 k が存在する.

$$\tilde{N}_2(t_{k-1}^{(2)}) = \tilde{N}_2^{np}(t_{k-1}^{(2)}) \quad , \quad \begin{cases} \tilde{N}_2(t_k^{(2)}) &= \tilde{N}_2(t_{k-1}^{(2)}) \\ \tilde{N}_2^{np}(t_k^{(2)}) &= \tilde{N}_2^{np}(t_{k-1}^{(2)}) + 1 \end{cases} \quad (8)$$

この式は時刻 $t_k^{(2)}$ において、グルーピングモデルではグルーピングが生じ、Non-priority モデルではグルーピングが生じていないことを表している。ここで、時刻 $t_k^{(2)}$ において、Non-priority モデルではグルーピングが生じていないことに着目すると、Non-priority モデルについて、次の二つの場合のどちらかが成り立つ。以降、それぞれの場合について考える。

Case1: 時刻 $t_{k-1}^{(2)}$ に到着したタイプ 2 の客はグルーピングが生じ、そのグループがちょうど N 人になった

Case2: 時刻 $t_{k-1}^{(2)}$ に到着したタイプ 2 の客は時刻 $t_k^{(2)}$ 以前に未完成グループのまま退去する

まず Case1 について考える。条件より $J^{np}(t_{k-1}^{(2)}) = 0, J^{np}(t_{k-2}^{(2)}) = N-1, \dots, J^{np}(t_{k-N}^{(2)}) = 1$ となるため、次の式が成り立つ。

$$\tilde{N}^{np}(t_{k-1}^{(2)}) = \tilde{N}^{np}(t_{k-1-j}^{(2)}) = \tilde{N}^{np}(t_{k-N}^{(2)}) = \tilde{N}^{np}(t_{k-N-1}^{(2)}) + 1 \quad (j = 1, \dots, N-1) \quad (9)$$

一方で、グルーピングモデルにおいて、時刻 $t_k^{(2)}$ に到着したタイプ 2 の客はグルーピングできていることから、 $J(t_{k-1}^{(2)}) = n \leq N-1$ と表すことができる。ただし、 n は自然数である。すると、 $J(t_{k-1}^{(2)}) = n, J(t_{k-2}^{(2)}) = n-1, \dots, J(t_{k-n}^{(2)}) = 1$ となるため、次の式が成り立つ。

$$\tilde{N}(t_{k-1}^{(2)}) = \tilde{N}(t_{k-n}^{(2)}) = \tilde{N}(t_{k-n-1}^{(2)}) + 1 \quad (10)$$

よって、(8), (9), (10) より、 $\tilde{N}(t_{k-n-1}^{(2)}) < \tilde{N}^{np}(t_{k-n-1}^{(2)})$ となるが、これは $n = k$ で初めて $\tilde{N}_2(t_n^{(2)}) < \tilde{N}_2^{np}(t_n^{(2)})$ となることに矛盾する。よって Case1 はあり得ないことが分かる。

次に Case2 について考える。補題 6.3 より、 $W^{np}(t_{k-1}^{(2)}) \geq W(t_{k-1}^{(2)})$ である。一方、グルーピングモデルでは(時刻 $t_k^{(2)}$ でグルーピングが生じているため) $[t_{k-1}^{(2)}, t_k^{(2)}]$ 間は待ち行列は空にならない。一方、Non-priority モデルでは未完成グループは行列の最後尾にいることから、未完成グループが時刻 $t_k^{(2)}$ までに退去することは、待ち行列が空になることを意味する。これは $W^{np}(t_{k-1}^{(2)}) \geq W(t_{k-1}^{(2)})$ であることと矛盾する。よって Case2 はあり得ないことが分かる。□

6.3. グルーピングモデルの安定条件

(1), (2), 定理 6.2 より、サンプルパスごとに $\rho \leq \rho^{kp}$ となるため、 K -priority モデルが安定状態であることは、グルーピングモデルが安定状態であることの十分条件となる。一方で、(1), (2), 定理 6.4 より、サンプルパスごとに $\rho \geq \rho^{np}$ となるため、Non-priority モデルが安定状態であることは、グルーピングモデルが安定状態であることの必要条件となる。ここで、特に K -priority モデルにおいて $K \rightarrow \infty$ の極限を考えると、

$$\lim_{K \rightarrow \infty} \rho^{kp} = \rho_1 + \lim_{K \rightarrow \infty} \frac{\rho_2}{N - (1 + \rho_2)^{-(K-1)} \sum_{j=1}^{N-1} \sum_{k=1}^j \left(\frac{\rho_2}{1 + \rho_2} \right)^{k-1} \binom{K+k-3}{k-1}} = \rho_1 + \frac{\rho_2}{N},$$

となり、 K -priority モデルの安定条件と Non-priority モデルの安定条件が一致することにより、グルーピングモデルが安定状態であることの必要条件と十分条件が一致するため、 $\rho_1 + \rho_2/N < 1$ はグルーピングモデルが安定状態であるための必要十分条件になる。

定理 6.5. 次の式を満たすとき、グルーピングモデルは安定状態をとる。

$$\rho_1 + \frac{\rho_2}{N} < 1.$$

参考文献

- [1] 塩田茂雄, 河西憲一, 豊泉洋, 会田雅樹, 待ち行列理論の基礎とその応用, 共立出版, 2014. 川島幸之助 (監修).

A. 補題 4.1 の証明

まず, $\tilde{\pi}_{i,j}$ を $\tilde{\pi}_0$ を用いて表すために, $\tilde{\pi}D^{(K)} = \mathbf{0}$ を以下のように書き下す.

$$\begin{aligned}\mu\tilde{\pi}_{i+1,1} &= (\mu + \lambda_2)\tilde{\pi}_{i,1}, \quad i = 1, \dots, K-2, \\ \lambda_2\tilde{\pi}_{i,j-1} + \mu\tilde{\pi}_{i+1,j} &= (\mu + \lambda_2)\tilde{\pi}_{i,j}, \quad i = 1, \dots, K-2, \quad j = 2, \dots, N-1 \\ \lambda_2\tilde{\pi}_0 &= (\mu + \lambda_2)\tilde{\pi}_{K-1,1}, \\ \lambda_2\tilde{\pi}_{K-1,j-1} &= (\mu + \lambda_2)\tilde{\pi}_{K-1,j}, \quad j = 2, \dots, N-1.\end{aligned}$$

上記の連立方程式を解くことにより, 以下が得られる.

$$\tilde{\pi}_{i,j} = a_{i,j}(1-q)^{K-1-i}q^j\tilde{\pi}_0, \quad q \stackrel{\text{def}}{=} \left(\frac{\rho_2}{1+\rho_2}\right)$$

ここで $a_{i,j}$ ($i = 1, \dots, K-1, \quad j = 1, \dots, N-1$) は以下を満たす定数である.

$$\begin{aligned}a_{i,j} &= \sum_{k=i}^{K-1} a_{k,j-1}, \quad i = 1, \dots, K-1, \quad j = 1, \dots, N-1 \\ a_{i,1} &= 1, \quad i = 1, \dots, K-1.\end{aligned}$$

これより $a_{i,j} = \binom{K-i+j-2}{j-1}$ が得られる. さて $g(j) \stackrel{\text{def}}{=} \sum_{i=1}^{K-1} \tilde{\pi}_{i,1}$ とおくと,

$$\begin{aligned}g(1) &= \sum_{i=1}^{K-1} \tilde{\pi}_{i,1} = q \sum_{i=1}^{K-1} a_{i,1}(1-q)^{K-1-i}\tilde{\pi}_0 \\ &= q \sum_{i=0}^{K-2} (1-q)^i\tilde{\pi}_0 = (1 - (1-q)^{K-1})\tilde{\pi}_0.\end{aligned}$$

また,

$$\begin{aligned}g(j) &= \sum_{i=1}^{K-1} \tilde{\pi}_{i,j} = q^j \sum_{i=1}^{K-1} a_{i,j}(1-q)^{K-1-i}\tilde{\pi}_0 = q^j \sum_{i=1}^{K-1} \sum_{k=i}^{K-1} a_{k,j-1}(1-q)^{K-1-i}\tilde{\pi}_0 \\ &= q^j \sum_{k=1}^{K-1} a_{k,j-1} \sum_{i=1}^k (1-q)^{K-1-i}\tilde{\pi}_0 = q^j \sum_{k=1}^{K-1} a_{k,j-1}(1-q)^{K-1-k} \sum_{i=1}^k (1-q)^i\tilde{\pi}_0 \\ &= q^{j-1} \sum_{k=1}^{K-1} a_{k,j-1}((1-q)^{K-1-k} - (1-q)^{K-1})\tilde{\pi}_0 \\ &= \sum_{i=1}^{K-1} a_{i,j-1}q^{j-1}(1-q)^{K-1-i}\tilde{\pi}_0 - q^{j-1}(1-q)^{K-1} \sum_{k=1}^{K-1} a_{k,j-1}\tilde{\pi}_0 \\ &= \sum_{i=1}^{K-1} \tilde{\pi}_{i,j-1} - q^{j-1}(1-q)^{K-1}a_{1,j}\tilde{\pi}_0 = g(j-1) - q^{j-1}(1-q)^{K-1}a_{1,j}\tilde{\pi}_0.\end{aligned}$$

つまり $g(j)$ についての漸化式が得られる. この漸化式を解くことにより

$$\begin{aligned}g(j) &= g(1) - q(1-q)^{K-1}a_{1,2}\tilde{\pi}_0 - \dots - q^{j-1}(1-q)^{K-1}a_{1,j}\tilde{\pi}_0 \\ &= \tilde{\pi}_0 - (1-q)^{K-1} \sum_{k=1}^j q^{k-1} \binom{K+k-3}{k-1} \tilde{\pi}_0.\end{aligned}$$

全確率が 1 であることより

$$\begin{aligned}1 &= \tilde{\pi}_0 + \sum_{j=1}^{N-1} \sum_{i=1}^{K-1} \tilde{\pi}_{i,j} = \tilde{\pi}_0 + \sum_{j=1}^{N-1} g(j) \\ &= \tilde{\pi}_0 \left(N - (1-q)^{K-1} \sum_{j=1}^{N-1} \sum_{k=1}^j q^{k-1} \binom{K+k-3}{k-1} \right).\end{aligned}$$

したがって

$$\tilde{\pi}_0 = \left(N - (1-q)^{K-1} \sum_{j=1}^{N-1} \sum_{k=1}^j q^{k-1} \binom{K+k-3}{k-1} \right)^{-1}.$$

B. 補題 6.1 の証明

グルーピングモデルのもとで、到着時に未完成グループがないタイプ 2 の客の到着時刻列を $(\tilde{t}_1^{(2)}, \tilde{t}_2^{(2)}, \dots)$ とする。さらに、 $(t_1^{(1)}, t_2^{(1)}, \dots)$ (タイプ 1 の客の到着時刻列) と $(\tilde{t}_1^{(2)}, \tilde{t}_2^{(2)}, \dots)$ を一緒にして、昇順に番号を振りなおしたものを $(t_1, t_2, \dots)$ とする。当初の $G_1+G_2/GI/1$ キューは到着時刻列が $(t_1, t_2, \dots)$ 、サービス時間列が $(\sigma_1, \sigma_2, \dots)$ の $G/GI/1$ キューと等価である。同様に、 K -priority モデルのもとで、到着時にペアリングが行われないタイプ 2 の客の到着時刻列とタイプ 1 の客の到着時刻列を一緒にして、昇順に番号を振りなおしたものを $(t_1^{kp}, t_2^{kp}, \dots)$ とする。(3) は $t_n \leq t < t_{n+1}$ かつ $t_n^{kp} \leq t < t_{n+1}^{kp}$ が満たされるならば

$$t_n \geq t_n^{kp}, \quad (11)$$

であることと等価である。時刻 t で客の到着がないならば、 $t_{l+1}^{kp} \stackrel{\text{def}}{=} t$ とおくとリンドレーの方程式より、

$$\begin{aligned} W^{kp}(t) &= \max \left\{ 0, \max_{k; 1 \leq k \leq n} \sum_{l=k}^n (\sigma_l - (t_{l+1}^{kp} - t_l^{kp})) \right\} \\ &= \max \left\{ 0, \max_{k; 1 \leq k \leq n} \left(t_n^{kp} - t + \sum_{l=n}^{k'-1} \sigma_l \right) \right\}, \end{aligned}$$

であること、同様に

$$W(t) = \max \left\{ 0, \max_{k; 1 \leq k \leq n} \left(t_n - t + \sum_{l=n}^{k'-1} \sigma_l \right) \right\},$$

であることに注意すると、(11) より $W^{kp}(t) \leq W(t)$ が成立する。時刻 t で客の到着がある場合も同様に成立する。

C. 補題 6.3 の証明

グルーピングモデルのもとで、到着時に未完成グループがないタイプ 2 の客の到着時刻列を $(\tilde{t}_1^{(2)}, \tilde{t}_2^{(2)}, \dots)$ とする。さらに、 $(t_1^{(1)}, t_2^{(1)}, \dots)$ (タイプ 1 の客の到着時刻列) と $(\tilde{t}_1^{(2)}, \tilde{t}_2^{(2)}, \dots)$ を一緒にして、昇順に番号を振りなおしたものを $(t_1, t_2, \dots)$ とする。当初の $G_1+G_2/GI/1$ キューは到着時刻列が $(t_1, t_2, \dots)$ 、サービス時間列が $(\sigma_1, \sigma_2, \dots)$ の $G/GI/1$ キューと等価である。同様に、Non-priority モデルのもとで、到着時にペアリングが行われないタイプ 2 の客の到着時刻列とタイプ 1 の客の到着時刻列を一緒にして、昇順に番号を振りなおしたものを $(t_1^{np}, t_2^{np}, \dots)$ とする。(7) は $t_n \leq t < t_{n+1}$ かつ $t_n^{np} \leq t < t_{n+1}^{np}$ が満たされるならば

$$t_n \leq t_n^{np}, \quad (12)$$

であることと等価である。時刻 t で客の到着がないならば、 $t_{l+1}^{np} \stackrel{\text{def}}{=} t$ とおくとリンドレーの方程式より、

$$\begin{aligned} W^{np}(t) &= \max \left\{ 0, \max_{k; 1 \leq k \leq n} \sum_{l=k}^n (\sigma_l - (t_{l+1}^{np} - t_l^{np})) \right\} \\ &= \max \left\{ 0, \max_{k; 1 \leq k \leq n} \left(t_n^{np} - t + \sum_{l=n}^{k'-1} \sigma_l \right) \right\}, \end{aligned}$$

であること、同様に

$$W(t) = \max \left\{ 0, \max_{k; 1 \leq k \leq n} \left(t_n - t + \sum_{l=n}^{k'-1} \sigma_l \right) \right\},$$

であることに注意すると、(12) より $W^{np}(t) \geq W(t)$ が成立する。時刻 t で客の到着がある場合も同様に成立する。

A downlink heterogeneous cellular network model with intra- and inter-tier correlated base station locations

小林 拓矢¹ 三好 直人² Harpreet S. Dhillon³
^{1,2} 東京工業大学 情報理工学院 ³ バージニア工科大学

概要： 本研究では基地局の間の様々な位置相関を考慮した HetNet モデルを提案する．具体的には，まず点同士の反発を表現できるジニブル点過程に従う点を配置し，それぞれの点に対して確率 $\alpha_m \in [0, 1]$ でマクロ基地局を，確率 $\alpha_c \in [0, 1]$ でその点を中心いくつかのsmall基地局を密集させて配置する．これらの確率やsmall基地局の強度関数を変化させることでマクロ基地局-マクロ基地局間やsmall基地局-small基地局間，マクロ基地局-small基地局間の位置相関をコントロールできる．

1. はじめに

近年のセルラネットワークは，トラフィックの急激な増加に対応するためにマクロ基地局中心のネットワークからいくつかの種類の小規模なsmall基地局を共存させることでより複雑で高密度なヘテロジニアスネットワーク (HetNet) へと変化している．HetNet の性能評価やシステム設計をするためには様々な基地局や無線端末の配置を考慮する必要がある．なぜならsmall基地局はマクロ基地局の電波を受信しにくい地点や無線端末を使用するユーザーが密集するホットスポット等へ意図的に多く配置される傾向があるためである．無線通信ネットワークの性能評価モデルとして，無線ノードの配置を空間点過程で表現する確率幾何モデルが注目されており，その多くは解析のしやすさのためそれぞれの配置を独立な定常ポアソン点過程でモデル化している．しかしそれらのモデルでは互いのノードの位置の相関を無視することになり，HetNet のような基地局や無線端末間の配置の相関を十分にとらえてるとは言えない．そのため，これらの相関を考慮した研究が近年盛んに行われている．例えば，それぞれの基地局が互いに反発する点の配置を表現できるジニブル点過程に従って配置された HetNet モデルの解析や [1]，ホットスポットに密集するユーザーとsmall基地局の組み合わせを同じクラスターセンターを持つポアソンクラスター過程で表現した HetNet モデルの解析が挙げられる [2]．

本研究では基地局の間の様々な位置相関を考慮した HetNet モデルを提案する．図 1 では 2 種類の基地局 (三角がマクロ基地局，丸がsmall基地局) の配置を構築する流れを表している．この図の例ではマクロ基地局同士が近づき過ぎないように配置しつつ，マクロ基地局から離れた地点にsmall基地局を配置できていることがわかる．さらに本研究ではこのモデルに対して二次元平面に一樣に分布するユーザーが受信する信号対干渉比 (SIR) の補分布 (被覆確率) を数値計算可能な形で導出し，基地局の強度やばらつき具合がネットワークに与える影響について考察する．

2. 提案モデル

提案モデルでは基地局の配置を以下の様に決める．まず 2 次元平面上に強度 λ の親点過程 $\Phi = \{X_i\}_{i \in \mathbb{N}}$ を点同士の反発を表現できるジニブル点過程で配置する (図 1 左)．それぞれの点に対して確率 $\alpha_m \in [0, 1]$ でマクロ基地局を割り当て，確率 $\alpha_c \in [0, 1]$ でsmall基地局のホットスポットを割り当てる (図 1 中央)．最後にホットスポットの点を中心にsmall基地局を有限なポアソン点過程に従って配置する (図 1 右)．それぞれの確率は $\alpha_m + \alpha_c \in (0, 1]$ を満たし，確率 $1 - \alpha_m - \alpha_c$ で Φ の点は除去される．それぞれのホットスポットに密集するsmall基地局の数は平均 μ のポアソン分布に従い，強度関数は $\lambda_s(x) = \mu f_s(|x|)$ で与えられるものとする (ただし， $\int_0^\infty f_s(u) u du = (2\pi)^{-1}$)．親の点同士は反発しているため，その強度 λ や α_m を調整することでマクロ基地局同士の反発をコントロールできる．またsmall基地局のホットスポットもジニブル点過程の点であるため，マクロ基地局-small基地局間の位置相関も表現できる．

3. 評価指標

$\bar{\Phi}_\xi = \{(Z_i, \xi_i)\}$ を基地局全体の配置と種類を表すマーク付き点過程とし， $\xi_i = m$ はマクロ基地局， $\xi_i = s$ でsmall基地局を表す．マクロ基地局とsmall基地局はそれぞれ一定の送信電力 $p_m, p_s > 0$ を持ち，距離 $r > 0$ に応じた電力の減衰を表すパスロス関数 $\ell_m(r), \ell_s(r)$ を持つ．レイリーフェージングを仮定し，基地局 (Z_i, ξ_i) の信号が受けるフェージングの影響 H_i はそれぞれ互いに独立な平均 1 の指数分布に従うとする．原点にいるユーザーが基地局 (Z_i, ξ_i) から受信する信号の強さは $p_{\xi_i} H_i \ell_{\xi_i}(|Z_i|)$ で与えられる．ユーザーはバイ

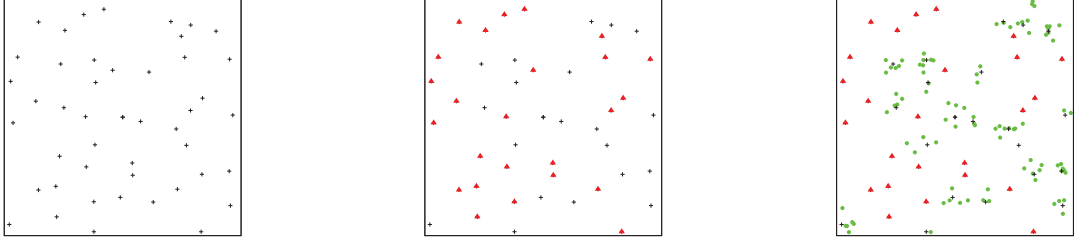


図 1: $\alpha_m = \alpha_c = 1/2$ の場合の 2-tier HetNet モデルのサンプルパス構築

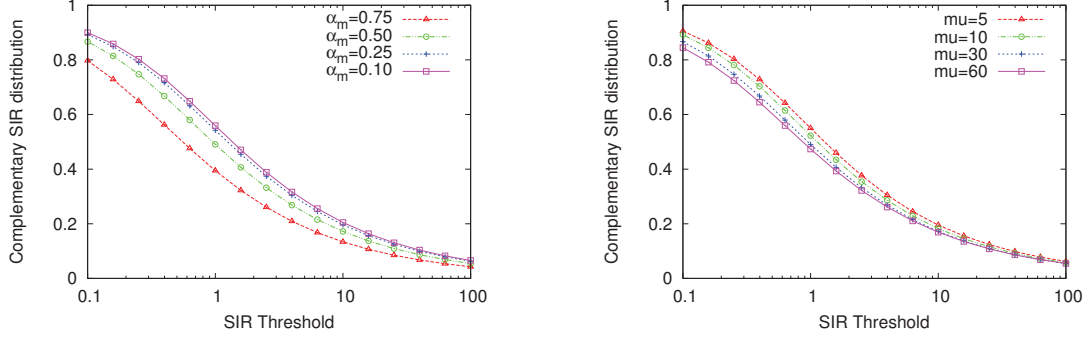


図 2: $\theta_m = \theta_s$ とし, α_m (左) と μ (右) を変化させた場合の 2-tier HetNet モデルの被覆確率

アス b_{ξ_i} を考慮した平均受信電力 $b_{\xi_i} p_{\xi_i} \ell_{\xi_i}(|Z_i|)$ が最大となる基地局と通信すると仮定し, 簡単のためその基地局を (Z_1, ξ_1) とする. このとき原点に位置するユーザーが受信する信号対干渉比 SIR_0 は以下で与えられる.

$$\text{SIR}_0 = \frac{p_{\xi_1} H_1 \ell_{\xi_1}(|Z_1|)}{\sum_{i=2}^{\infty} p_{\xi_i} H_i \ell_{\xi_i}(|Z_i|)}.$$

本研究では SIR_0 が閾値 θ_{ξ_1} を超える確率 (被覆確率) を数値計算可能な形で導出する. 解析ではまず親点過程 Φ を条件づけた下で通信する基地局の種類ごとに被覆確率を与え, それらの確率が距離 $\{|X_i|\}_{i \in \mathbb{N}}$ のみに依存することと, Φ がジニブル点過程の場合にそれらの距離が互いに独立なガンマ分布に従うことを利用している.

4. 主結果

定理 1. 提案モデルにおける被覆確率は以下で与えられる.

$$\begin{aligned} \mathbb{P}(\text{SIR}_0 > \theta_{\xi_1}) &= \int_0^\infty \sum_{i \in \mathbb{N}} f_i(t | \pi\lambda) \left[\alpha_m \prod_{j \in \mathbb{N} \setminus \{i\}} \left(\int_0^\infty f_j(s | \pi\lambda) v_m(\sqrt{t}, \sqrt{s}) ds \right) \right. \\ &\quad \left. + \mu \alpha_c \int_0^\infty g(r | \sqrt{t}) (1 - B_{s,s}(r, \sqrt{t})) \prod_{j \in \mathbb{N} \setminus \{i\}} \left(\int_0^\infty f_j(s | \pi\lambda) v_s(r, \sqrt{s}) ds \right) dr \right] dt. \end{aligned}$$

ただし, $v_k(x, y) = 1 - \alpha_m A_{k,m}(x, y) - \alpha_c B_{k,s}(x, y)$, $f_i(t | \pi\lambda) = (\pi\lambda)^i t^{i-1} e^{-\pi\lambda t} / (i-1)!$,

$$A_{k,l}(x, y) = 1 - \mathbf{1}_{\{y > r_{k,l}(x)\}} \left(1 + \theta_k \frac{b_k \ell_l(y)}{b_l \ell_l(r_{k,l}(x))} \right)^{-1}, \quad r_{k,l}(x) = \ell_l^{-1}(b_k p_k \ell_k(x) / (b_l p_l)),$$

$$B_{k,l}(x, y) = 1 - \exp \left(-\mu \int_0^\infty A_{k,l}(x, u) g(u | y) du \right), \quad k, l \in \{m, s\}, x, y \geq 0,$$

$$g(u | y) = 2u \int_0^\pi f_s(\sqrt{u^2 + y^2 - 2uy \cos \phi}) d\phi, \quad u, y \geq 0.$$

参考文献

- [1] I. Nakata and N. Miyoshi, “Spatial stochastic models for analysis of heterogeneous cellular networks with repulsively deployed base stations,” *Performance Evaluation*, vol. 78, pp. 7–17, 2014.
- [2] C. Saha, H. S. Dhillon, N. Miyoshi and J. G. Andrews, “Unified analysis of HetNets using Poisson cluster process under max-power association,” *arXiv:1812.01830*, 2018.

A Markov Model of Eusociality at its Origin

Hiroshi Toyoizumi
Waseda University

Jeremy Field
University of Exeter

Abstract To understand advantages and disadvantages of various eusocial strategies, we use a transient Markov arrival process and branching process modeling the group dynamics in a cooperatively-breeding nest and its offspring dispersals. We show how to derive the basic reproductive number and the extinction probability, and give some insight to the origin of eusociality.

1. Introduction

The origin and maintenance of eusociality is a question of major interest in evolutionary biology. In eusocial and cooperatively-breeding taxa, some individuals, known as helpers or workers, forfeit their own reproduction to aid the reproduction of other individuals known as queens or breeders. In a recent theoretical paper, Fu et al. (2015) took a new approach by using Markov chain and branching processes to model colony and population dynamics respectively. We here extend the model of Fu et al. (2015) to include:

1. Worker mortality: we model not only queen but also workers' death rates.
2. Inheritance of nest: instead of queen death causing catastrophic group failure, we assume that if the queen dies a surviving worker can take over as the egg-layer so that the group continues to produce offspring.
3. Productivity related linearly to group size: rather than productivity increasing only above a threshold group size, we assume that each worker increases productivity by the same amount.

We use branching process (see Durrett (2012) for example) and transient Markovian arrival process (Latouche et al. 2003) to analyze our models and to derive both the basic reproductive number and the extinction probability. We find that allowing inheritance is particularly important for understanding the origin of eusociality. Note that Toyoizumi and Field (2014) uses a similar technique using quasi-birth and death process but only to derive the basic reproductive number.

2. Solitary Model

We start with the basic solitary model, which introduces the concepts we will use. Like Fu et al. (2015), we assume that a solitary lineage starts with a single foundress that builds a nest and starts producing offspring. Birth intervals and the lifespan of the foundress are independent and exponentially distributed with b_0 being the offspring birth rate and d_0 being the death rate of the foundress. Let X be the number offspring that the foundress produces in her life span, then we have

$$P(X = n) = \frac{d_0}{b_0 + d_0} \left(\frac{b_0}{b_0 + d_0} \right)^n. \quad (1)$$

This can be interpreted as the following stochastically equivalent setting: set neutral random events in time with the rate $b_0 + d_0$, where each event is birth with the probability $b_0/(b_0 + d_0)$ or death with the probability $d_0/(b_0 + d_0)$, so that the number of births until foundress death follows a geometric distribution as in (1).

Using the expectation of X , we can find the basic reproductive number R as

$$R = E[X] = \sum_{n=0}^{\infty} n \frac{d_0}{b_0 + d_0} \left(\frac{b_0}{b_0 + d_0} \right)^n = \frac{b_0}{d_0},$$

which is simply the product of the average life span $1/d_0$ and the birth rate b_0 . When R is greater than 1, one foundress produces more than one offspring, and the population grows exponentially.

In the solitary model, all offspring act as foundresses, start their own nests and begin reproducing immediately. The population is extinct when all foundresses are dead. We also investigate the populations extinction probability p , which can be analyzed by the probability generating function $\phi(z) = E[z^X]$ for $z \in [0, 1]$. Assuming that each offspring is marked red independently with the probability z , then $\phi(z)$ can be interpreted as the probability that all offspring are happened to be red.

The extinction probability p is known to be the minimal solution of the equation for the branching process (e.g. see Durrett (2012) for an example):

$$p = \phi(p) = E[p^X]. \quad (2)$$

Equation (2) can be interpreted as follows: given that the initial nest produces X independent dispersing foundresses, all of these X lineages go extinct with probability p^X , and thus the extinction probability p is simply the expected probability $E[p^X]$. By (1), given $R = b_0/d_0 > 1$, we have

$$\phi(z) = E[z^X] = \sum_{k=0}^{\infty} z^k \frac{d_0}{b_0 + d_0} \left(\frac{b_0}{b_0 + d_0} \right)^k = \frac{d_0}{b_0 + d_0 - b_0 z}.$$

Thus, the extinction probability p has to satisfy

$$0 = p(b_0 + d_0 - b_0 p) - d_0 = (-b_0 p + d_0)(p - 1),$$

and it is easy to show that $p = d_0/b_0 = 1/R$ is its minimal solution. Thus, the extinction probability and the basic reproductive number are essentially the same in the solitary model, with a larger reproductive number R always giving a smaller extinction probability p . However, we will see that there is not always such a simple relationship between extinction probability and the basic reproductive number.

3. Transient Markov Model for Estimating Dispersals

We can extend the above argument to more general models including group dynamics within a eusocial group. We use transient Markovian Arrival Process (MAP) (Latouche et al. 2003) for modeling the number of dispersals.

Consider an initial eusocial foundress that builds a nest and starts producing offspring. Under the eusocial strategy, some of her offspring may remain in the natal group as helpers, and we let $L(t)$ be the group size (i.e. the number of individuals in the group at time t including the foundress if she is still alive). We let X be the number of dispersing offspring (new foundresses) produced before the group termination time T that occurs either when no individual remains alive ($L(T) = 0$) or when a sudden catastrophic event happens to the group. We assume the group has a maximum size M .

For simplicity, we assume that the group dynamics of $L(t)$ can be analyzed by a Markov chain, which means that we can focus on only the current group size $L(t)$ to predict the future dynamics (see Figure 1). Suppose that group size $L(t) = i$. The group size increases with rate λ_i and decreases with rate μ_i . In addition, the nest may be terminated with rate γ_i due to some catastrophe (the queen's death in the model of Fu et al. (2015), or in our bet-hedging model of Section 6 perhaps due to attack by predators). When $L(t) = 1$, the death of the last remaining individual terminates the nest so that there can be no downward transition from $L(t) = 1$, and we set $\gamma_1 > 0$ and $\mu_1 = 0$. An offspring dispersal event occurs with rate ν_i , and each dispersed offspring is assumed to be a foundress that starts her own nest.

Let $D = (D(i, j))$ be the transition rate matrix of a general M -state Markov chain of $L(t)$, where the component $D(i, j)$ represents the transition rate from group size i to group size j . We decompose D by the transition with and without dispersal. Let D_0 be the transition rate matrix

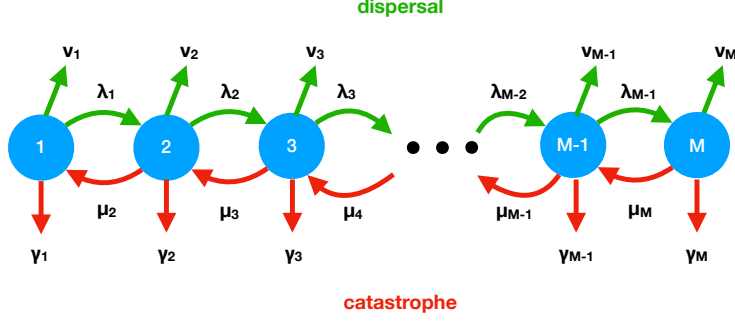


Figure 1: Transient Markov chain model of dispersal. Numbers in blue circles represent group size $L(t)$ (including the queen), which changes at the rates shown. Green arrows are related to the birth and red arrows are related to the death of individuals in the group or (downward-pointing red arrows) to catastrophic group failure. Note that upward-pointing green arrows indicating dispersal do not change the group size.

without dispersal, D_1 be the transition rate matrix with dispersal, and C be the transition rate matrix to catastrophe. To give an example, D_0 for Figure 1 is a tridiagonal matrix:

$$D_0 = \begin{pmatrix} -(\lambda_1 + \gamma_1 + \nu_1) & \lambda_1 & & & 0 \\ \mu_2 & -(\lambda_2 + \gamma_2 + \mu_2 + \nu_2) & \lambda_2 & & \\ & \mu_3 & -(\lambda_3 + \gamma_3 + \mu_3 + \nu_3) & \lambda_3 & \\ & & & \ddots & \\ 0 & & & \mu_M & -(\gamma_M + \mu_M + \nu_M) \end{pmatrix}$$

and D_1 is a diagonal matrix with $D_1 = \text{diag}(\nu_1, \nu_2, \dots, \nu_M)$ and C is also a diagonal matrix with $C = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_M)$.

Since there can be a catastrophic event, $L(t)$ is transient, and both $D = D_0 + D_1$ and D_0 are non-singular, where we assume that a catastrophe and a dispersal do not occur at the same time. Note that we could potentially apply this analysis to more general models than Figure 1, such as when two or more births can occur at the same time.

The transient probability distribution $\pi_i(t) = P(L(t) = i, \text{no catastrophe during } (0, t))$ is characterized by the differential equation $\frac{d}{dt}\pi(t) = \pi(t)D$ and its solution:

$$\pi(t) = \pi(0)e^{Dt},$$

where $\pi(t) = (\pi_1(t), \dots, \pi_M(t))$ and $e^{Dt} = \sum_{k=0}^{\infty} (Dt)^k / k!$. In general, e^{Dt} is called the semi-group operator of the system that can advance the clock of the model. Since the Markov chain generated by D stops (terminates) at the catastrophic event, $\sum_{i=1}^M \pi_i(t)$ may not attain 1. The missing quantity $1 - \sum_{i=1}^M \pi_i(t)$ represents the probability that the nest has already terminated by the time t .

Let T_i be the time duration that the group size is i , then its expectation can be obtained by

$$E[T_i] = E \left[\int_0^{\infty} 1_{\{L(t)=i\}} dt \right] = \int_0^{\infty} \pi_i(t) dt = (\pi(0)(-D)^{-1})_i,$$

which is equivalent to the lifespan $1/d_0$ in the solitary model.

Since the Markov chain generated by D_0 represents the process stopped either at the catastrophe or dispersal, similar arguments give its semi-group operator as

$$P(L(t) = j, \text{no dispersal and no catastrophe during } (0, t) | L(0) = i) = (e^{D_0 t})_{i,j}.$$

Let $F = \int_0^{\infty} e^{D_0 t} D_1 dt = (-D_0)^{-1} D_1$, which can be interpreted as the transition probability matrix on the successive events of dispersal. Note that F is a sub-stochastic matrix (at least one of the

sums of each row is less than 1) since there is the potential for a catastrophic event. Similarly, we define $G = (-D_0)^{-1}C$ which is the transition probability matrix from dispersal to catastrophe. The matrices F and G corresponds to $b_0/(b_0 + d_0)$ and $d_0/(b_0 + d_0)$ in the solitary model. Then,

$$(P(X = n, L(T-) = 1), \dots, P(X = n, L(T-) = M)) = \pi(0)F^n G. \quad (3)$$

where $\pi(0) = (1, 0, \dots, 0)$. The right-hand side of (3) can be regarded as starting from the initial distribution $\pi(0)$, repeating the dispersal transition F with n times and then a catastrophe transition G occurs. It is not difficult to see that

$$P(X = n) = \pi(0)F^n G\mathbf{1},$$

where $\mathbf{1} = (1, 1, \dots, 1)^T$. This is indeed equivalent to Theorem 4.1 in Latouche et al. (2003) with $f_0 = G\mathbf{1}$ and $f_1 = 0$ for the case of no catastrophe occurring at the same time as another event. Since F is sub-stochastic matrix, we have the identity $(I - zF) \sum_{n=0}^{\infty} z^n F^n = I$. Thus, we have

$$E[z^X] = \sum_{n=0}^{\infty} z^n P(X = n) = \pi(0)(I - zF)^{-1}G\mathbf{1}.$$

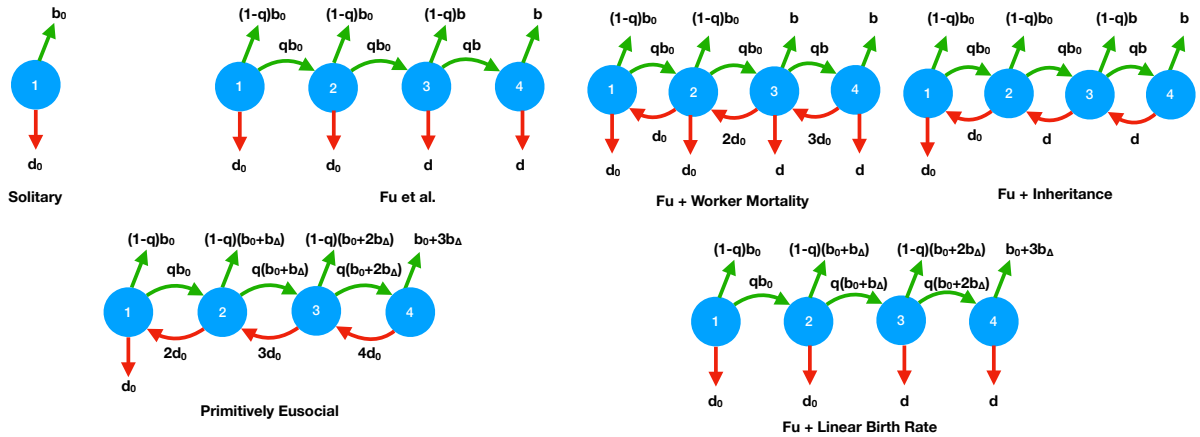
Also, since $(I - F) \sum_{n=0}^{\infty} nF^n = \sum_{n=1}^{\infty} F^n$, the basic reproductive number is obtained by

$$R = E[X] = \pi(0)F(I - F)^{-2}G\mathbf{1}. \quad (4)$$

Note that these expressions are the direct extension of the solitary case. Using these expressions, we can easily derive the basic reproductive number R and the extinction probability p in general models such as the one shown in Figure 1, especially for a small M .

4. Specific Eusocial Models

Using the general model described in Section 3, we now define the specific models shown schematically in Figure 2, including our Primitively Eusocial model.



(a) Solitary, Fu et al. eusocial and Prim Eusocial

(b) Auxiliary Models

Figure 2: Models Analysed (with $m = 3$ and $M = 4$).

First, we define the model used in Fu et al. (2015), which has: (1) no worker mortality (2) no inheritance following queen death, and (3) a threshold-type offspring birth rate, and a threshold-type queen death rate which causes a catastrophic event. In contrast with the solitary model, newly-born offspring now remain in the natal group with probability q (the staying ratio), where

they may help their mother. All newly-born offspring must leave the nest when $L(t) = M$. If group size reaches a threshold ($L(t) \geq m$), both the birth rate and the death rate change (positive effect of helpers). When the foundress dies, the group is terminated. Thus, in terms of our general model in Section 3, the Eusocial strategy of Fu et al. is described as:

$$\begin{aligned} \lambda_i &= \begin{cases} qb_0 & \text{for } 1 \leq i < m, \\ qb & \text{for } m \leq i \leq M-1, \\ 0 & \text{for } i = M. \end{cases} & \nu_i &= \begin{cases} (1-q)b_0 & \text{for } 1 \leq i < m, \\ (1-q)b & \text{for } m \leq i \leq M-1, \\ b & \text{for } i = M. \end{cases} \\ \gamma_i &= \begin{cases} d_0 & \text{for } 1 \leq i < m, \\ d & \text{for } m \leq i \leq M, \end{cases} \end{aligned} \quad (5)$$

and $\mu_i = 0$ for all i (there is no mortality except catastrophic group failure following the death of the queen). Using this approach, we successfully reproduced the results shown in Fu et al. (2015) Figures 2 and 3. Note that either simple probabilistic or linear algebra arguments on (4) lead to $R = \sum_{k=1}^M \left(\prod_{j=1}^{k-1} \frac{\lambda_j}{\gamma_j + \lambda_j} \right) \frac{\nu_k}{\gamma_k + \lambda_k}$, which coincides with (16) in Fu et al. (2015).

We next specify our Primitively Eusocial model (Figure 2a). As in Fu et al.'s model, offspring stay in their natal groups with probability q , but in contrast with the threshold-type offspring birth rate in Fu et al.'s model, so that workers contribute something even at small group sizes (linear birth rate). We also allow worker mortality at a rate proportional to group size, reflecting the cost of the linear increase in productivity. In contrast with Fu et al., when the foundress (queen) in the group dies, we allow inheritance of the egg-laying position: the group is not terminated and continues producing offspring (initially at the rate appropriate for the smaller group size) until the last individual in the group dies. Thus, our Primitively Eusocial strategy is described as:

$$\begin{aligned} \lambda_i &= \begin{cases} q\{b_0 + (i-1)b_\Delta\} & \text{for } 1 \leq i < M \\ 0 & \text{for } i = M. \end{cases} & \nu_i &= \begin{cases} (1-q)\{b_0 + (i-1)b_\Delta\} & \text{for } 1 \leq i < M \\ b_0 + (M-1)b_\Delta & \text{for } i = M. \end{cases} \\ \gamma_i &= \begin{cases} d_0 & \text{for } i = 1, \\ 0 & \text{for } 2 \leq i \leq M, \end{cases} & \mu_i &= \begin{cases} 0 & \text{for } i = 1, \\ d_0 i & \text{for } 2 \leq i \leq M, \end{cases} \end{aligned}$$

where b_Δ represents the productivity boost provided by each additional worker. In this case, by (4) or by the simple probabilistic argument based on the memoryless of exponential distribution, we can derive $R = \sum_{k=1}^M \left(\prod_{j=1}^{k-1} \frac{\lambda_j}{\mu_j} \right) \frac{\nu_k}{\mu_k}$ where we set $\mu_1 = \gamma_1$ for simplicity.

To separately examine the effect of each of our modifications to Fu et al.'s Eusocial strategy model (worker mortality, inheritance, linear birth rate), we use three auxiliary models; (1) Fu + worker mortality, (2) Fu + inheritance, and (3) Fu + linear birth rate (see Figure 2b). Each represents a single component of our Primitively Eusocial model.

In the Fu + worker mortality model, with each additional worker we add d_0 to the downward group size transition rate. Thus, in (5), we add the downward transition rate as

$$\mu_i = (i-1)d_0,$$

where the factor -1 reflects the fact that foundress death results in whole group failure and counted in the catastrophe rate γ_i .

In the Fu + inheritance model, we modify Fu et al.'s Eusocial strategy by assuming that one of the remaining workers inherits the egg-laying position when the queen dies, so that there is no whole group catastrophe rate for $L(t) \geq 2$, and we set:

$$\begin{aligned} \gamma_i &= \begin{cases} d_0 & \text{for } i = 1, \\ 0 & \text{for } i \geq 2, \end{cases} & \mu_i &= \begin{cases} d_0 & \text{for } 2 \leq i < m, \\ d & \text{for } m \leq i \leq M, \end{cases} \end{aligned}$$

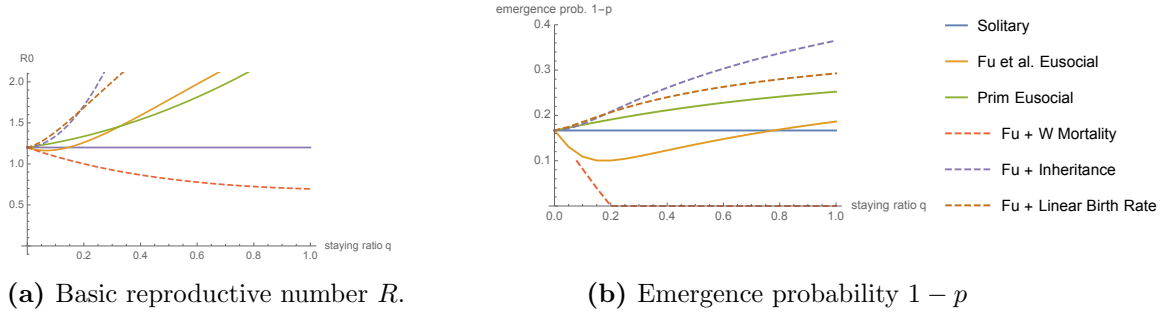


Figure 3: Comparison of Models in terms of R and in terms of the probability of avoiding extinction (emergence probability). The parameters are set by $M = m = 3$; $b_0 = 0.12$, $d_0 = 0.1$; $b = 0.45$, $d = 0.05$; $b_\Delta = (b - b_0)/(m - 1) = 0.165$.

Note that this model is otherwise specified as in (4), i.e. we do not allow worker deaths.

Finally, in the Fu + linear birth rate model, we assume that group productivity increases linearly with the addition of each worker, so that:

$$\lambda_i = \begin{cases} q\{b_0 + (i - 1)b_\Delta\} & \text{for } 1 \leq i < M \\ 0 & \text{for } i = M. \end{cases} \quad \nu_i = \begin{cases} (1 - q)\{b_0 + (i - 1)b_\Delta\} & \text{for } 1 \leq i < M \\ b_0 + (M - 1)b_\Delta & \text{for } i = M. \end{cases}$$

Otherwise, the model is specified as in (4). If we add all three modifications simultaneously, we obtain our full Primitively Eusocial model.

5. Numerical Examples

Numerical example derivations of the basic reproductive number R and the emergence probability $1 - p$ (probability of avoiding extinction) are shown in Figure 3. The maximum group size is assumed to be small ($m = M = 3$), which seems more realistic, when modelling the origin of eusociality, than the assumption ($m = 3$, $M = 100$) in Fu et al. (2015). In our Primitively Eusocial model and the Fu + linear birth rate model, the increase in productivity with group size is assumed to be a linear interpolation between b_0 and b ; thus we set $b_\Delta = (b - b_0)/(m - 1)$. Other parameters are identical to those in the middle graphs in Figure 2 of Fu et al. (2015).

In some of their numerical examples, Fu et al. (2015) found that both R and $1 - p$ decreased at smaller staying ratios q (see the orange lines in Figure 3), which they interpreted as reflecting a risk inherent in eusociality: the opportunity cost of worker production delaying the production of offspring that immediately disperse. However, in our Primitively Eusocial model, both R and $1 - p$ increase with q , so that this risk apparently disappears. We check what causes this difference by analysing the three auxiliary (partially modified) models; (1) Fu + worker mortality, (2) Fu + inheritance, and (3) Fu + linear birth rate (see Figures 3). Adding worker mortality is a big disadvantage for eusociality, while adding either inheritance or a linear birth rate are advantages (see Figure 3), and the sum of the latter two positive effects is larger than the negative effect of worker mortality, so that the risk disappears in our overall Primitively Eusocial model.

For a detailed comparison of the models (Figure 4), we illustrate the number of dispersing offspring produced at each group size $L(t)$ when the staying ratio is small ($q = 0.1$) and large ($q = 0.7$), which are easily estimated by the vector $\pi(0)(-D)^{-1}D_1$. Also, we include the expected time duration at each size. First, consider the case when the staying ratio is small ($q = 0.1$). Because we allow worker mortality, group size reaches the maximum ($M = 3$) less frequently in the Primitively Eusocial model than in Fu et al.'s eusocial strategy, so that productivity at this group size is smaller (see also Fu + worker mortality). However, inheritance results in greater productivity at $L(t) = 1$ than in Fu et al.'s model (the nest always returns to group size 1 before group termination), and this is the main source of the improved performance of eusociality in the Primitively Eusocial model (and the Fu + inheritance model). Inheritance is important because it

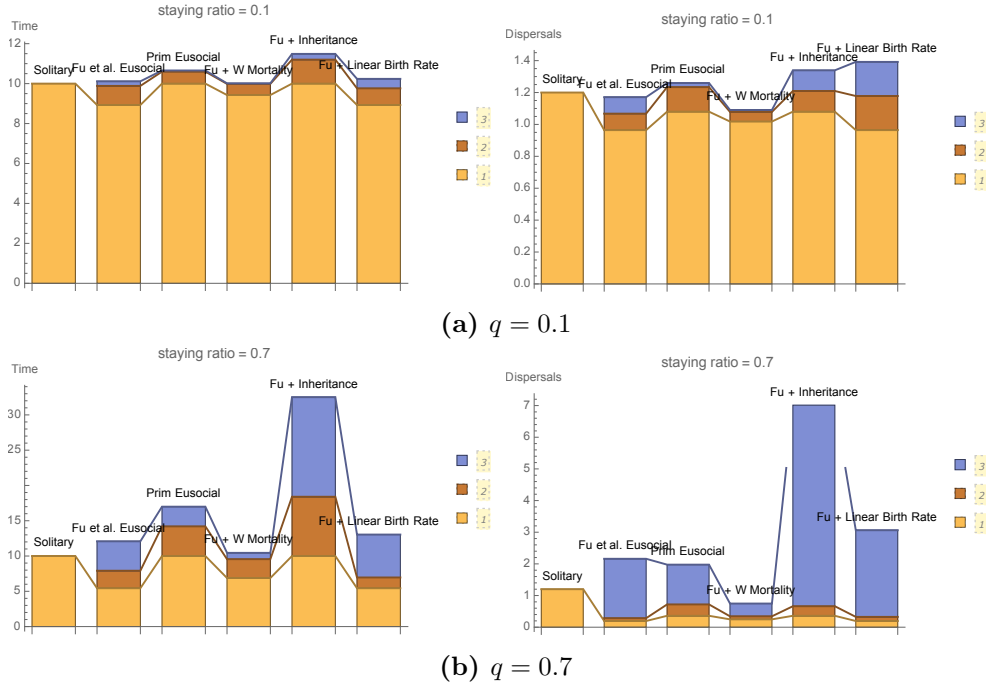


Figure 4: Mean number of dispersals and mean time spent at each group size, where colours represent different group sizes ($q = 0.1$ or 0.7 ; $M = m = 3$; $b_0 = 0.12, d_0 = 0.1$; $b = 0.45, d = 0.05$; $b_\Delta = (b - b_0)/(m - 1) = 0.165$)

avoids the wasted investment in worker production that occurs when catastrophic group termination follows queen death in Fu et al.'s eusocial model. Instead allowing productivity to increase linearly with group size does lead to increased dispersal at group size $L(t) = 2$, but it is a smaller effect (compare the total dispersals at $L(t) = 1$ and 2 of Fu + inheritance and Fu + linear birth rate). Next, consider the case when the staying ratio is high ($q = 0.7$). Fu et al.'s eusocial strategy now spends longer at the maximum group size thanks to the low queen death rate. This advantage is significant because reaching to the maximum ($M = 3$) is relatively easy now. At high staying ratios, Fu et al.'s eusocial strategy can therefore perform better in terms of the basic reproductive number $E[X]$ for some combinations of parameter values, as shown in Figure 3a.

To obtain an idea of the robustness of our results when maximum group size is larger, we also illustrate the number of dispersals when $m = 3, M = 10$ (see Figure 5). The results are qualitatively similar to those for $m = 3, M = 3$, but the emergence probability for the eusocial strategy of Fu et al. now continuously decreases with increasing q . The effect of allowing just inheritance is much larger than in the case of $m = 3, M = 3$, because more offspring can potentially become helpers when $M = 10$, so that more investment is wasted in Fu et al.'s eusocial strategy when catastrophic group failure occurs.

Further, we checked the robustness of our result by looking at different b/d ratio in Figure 6, which clearly shows that the primitively eusocial model performs better especially at smaller staying ratios q .

6. Bet-Hedging

We also view our general model in terms of spatial bet-hedging (Starrfelt and Kokko 2012), where the solitary strategy effectively spreads risk by producing offspring that disperse individually as opposed to staying on the natal nest under the eusocial strategy. To illustrate this, we consider a simple case of the Primitively Eusocial model with $M = 2$ and $q = 1$ (all offspring become helpers below the maximum group size). We additionally specify completely linear productivity $b_\Delta = b_0$ and a rate of catastrophic group failure c that is independent of group size (see Figure 7):

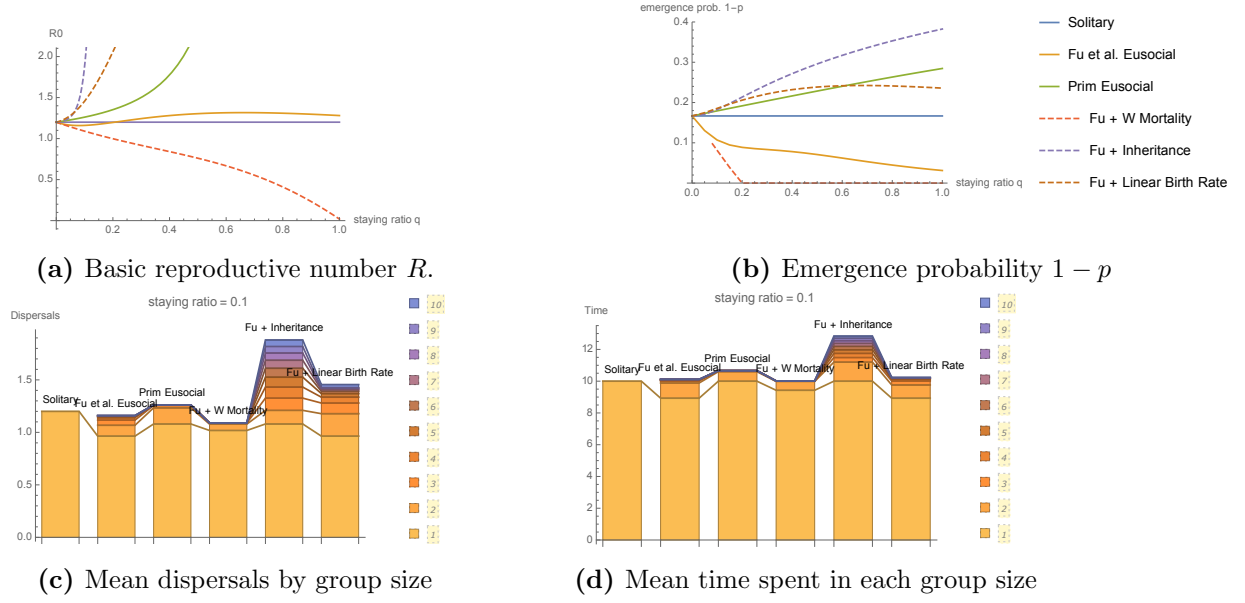


Figure 5: Comparison of Models as in Fig. 3 and 4a, but with $M = 10$. The parameters are set by $m = 3, M = 10; b_0 = 0.12, d_0 = 0.1; b = 0.45, d = 0.05; b_\Delta = (b - b_0)/(m - 1) = 0.165$.

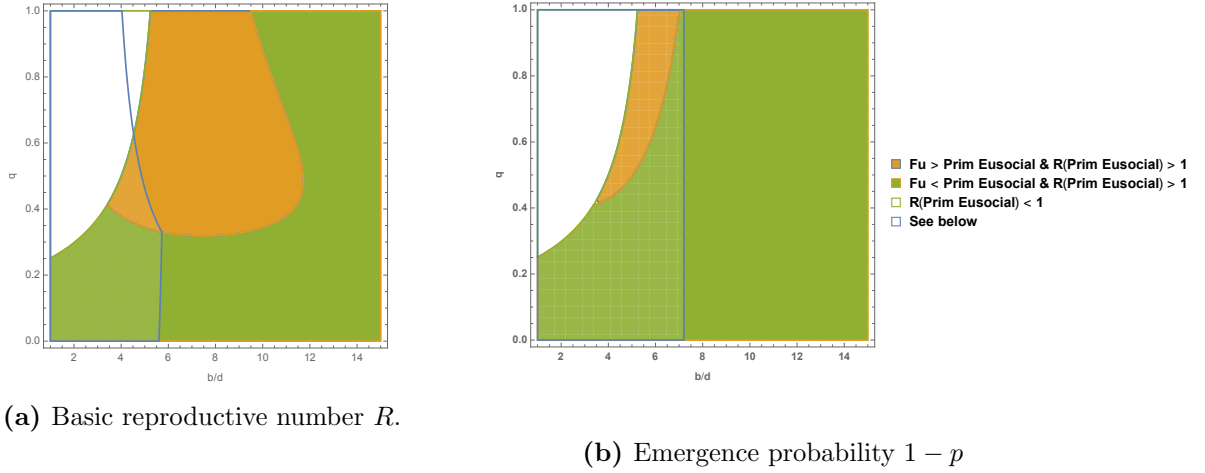


Figure 6: Region graphs comparing Fu et al.'s solitary and eusocial strategies with our Primitively Eusocial strategy in terms of basic reproductive number and emergence probability in the $b/d - q$ plane. Other parameter values are $m = 3, M = 3; b_0 = 0.12, d_0 = 0.1; d = 0.05; b_\Delta = (b - b_0)/(m - 1)$. The white region indicates the area where the Primitively Eusocial strategy is certain to go extinct (its basic reproductive number is $R = E[X] < 1$). The area surrounded by the blue boundary is where the solitary strategy performs best so that in (b), for example, our solitary strategy performs best at all q values when b/d is low, whereas Primitive Eusociality performs best when b/d is high.

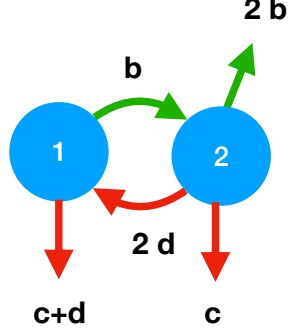


Figure 7: Schematic for a simple eusocial model to illustrate one form of bet-hedging

$$D_0 = \begin{pmatrix} -(d+b+c) & b \\ 2d & -(2d+2b+c) \end{pmatrix}, D_1 = \begin{pmatrix} 0 & 0 \\ 0 & 2b \end{pmatrix}, C = \begin{pmatrix} c+d & 0 \\ 0 & c \end{pmatrix}.$$

Then, with some direct matrix calculations as in Section 3, the probability generating function for the number of dispersals X in this model is obtained by

$$E[z^X] = \frac{bc + (c+d)(c+2d) + 2b(c+d)(1-z)}{bc + (c+d)(c+2d) + 2b(c+d+b)(1-z)}, \quad (6)$$

and also the basic reproductive number is obtained by

$$R_{\text{eusocial}} = E[X] = \frac{2b^2}{bc + (c+d)(c+2d)}. \quad (7)$$

We compare these results with the solitary model under the same catastrophe rate c .

Before comparing the two strategies, however, we check the results in the case with a zero catastrophe rate: these should be the same for solitary and eusocial. As described in Section 2, given $b > d$, we have the extinction probability $p = d/b$ and the basic reproductive number $E[X] = b/d$ for the solitary model. If we set $c = 0$ in (6), we have

$$E[z^X] = \frac{2d^2 + 2bd(1-z)}{2d^2 + 2b(b+d)(1-z)},$$

and set $z = d/b$ to have

$$E[z^X]|_{z=d/b} = \frac{2bd^2 + 2bd(b-d)}{2bd^2 + 2b(b+d)(b-d)} = \frac{d}{b}.$$

This shows that the eusocial model has the same extinction probability as solitary model as predicted. However, if we set $c = 0$ in (7), we have

$$R_{\text{eusocial}} = \left(\frac{b}{d}\right)^2 > \frac{b}{d} = R_{\text{solitary}},$$

thus the eusocial model has always a larger basic reproductive number. This is because the lifespans of the two types of group are different, and suggests that generally, comparing the strategies in terms of the basic reproductive number could be misleading. Using the extinction probability is preferable.

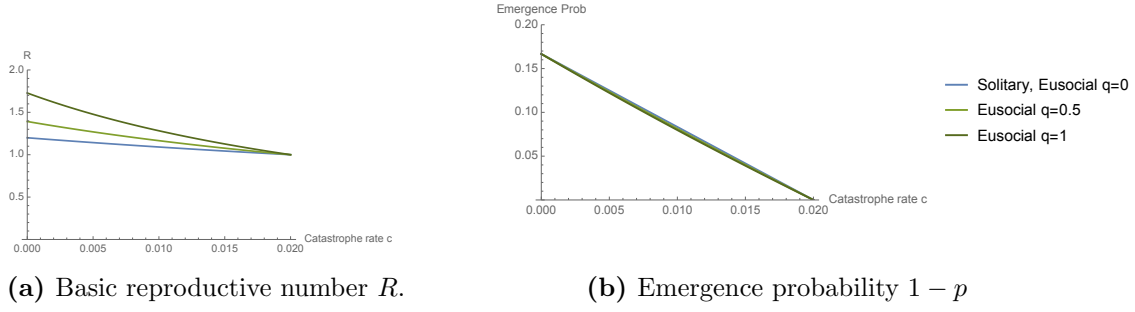


Figure 8: An example where the solitary strategy has an advantage through risk-spreading. The parameters are set by $M = 3; b = 0.12, d = 0.1$.

Next, we compare the two strategies including a positive catastrophe rate. Given $b > c + d$, we have the extinction probability $p = (d + c)/b$ and the basic reproductive number $E[X] = b/(c + d)$ for the solitary model with catastrophe. If $b > c + d$, the simple eusocial model has

$$R_{\text{eusocial}} = \frac{2b^2}{bc + (c + d)(c + 2d)} > \frac{2b^2}{bc + b(c + 2d)} = \frac{b}{c + d} = R_{\text{solitary}},$$

thus the eusocial model has always a larger basic reproductive number, but this may not be meaningful as just discussed. For the extinction probability, taking $z = (d + c)/b$ in (6), then we have

$$\begin{aligned} E[z^X]_{|z=(d+c)/b} &= \frac{bc + (c + d)(c + 2d) + 2(c + d)(b - c - d)}{bc + (c + d)(c + 2d) + 2(b + c + d)(b - c - d)} \\ &= \frac{bc + (c + d)(2b - c)}{c(b - c - d) + 2b^2}. \end{aligned}$$

Now given $b > d + c$, we have

$$E[z^X]_{|z=(d+c)/b} - \frac{d + c}{b} = \frac{c(b - c - d)^2}{b \{2b^2 + c(b - c - d)\}} > 0.$$

Since $E[z^X]$ is a decreasing function of z , the extinction probability satisfying $p = E[p^X]$ is greater than $(d + c)/p$. Thus, the extinction probability for the simple eusocial model is always larger than for the solitary model. The solitary strategy performs better if we have a positive catastrophe rate, and the only difference between the two strategies is the fact that offspring produced under the solitary strategy initiate new nests.

Figure 8 shows the basic reproductive number and the extinction probability for the case of $M = 3$. We can see that as shown above for the case when $M = 2$, the solitary strategy performs slightly better than eusociality in terms of the emergence probability. However, the advantage is marginal, and so may be overcome by only a slightly super-linear productivity for the eusocial strategy.

References

- Richard Durrett. *Essentials of stochastic processes*. Springer Science & Business Media, 2012.
- Feng Fu, Sarah D Kocher, and Martin A Nowak. The risk-return trade-off between solitary and eusocial reproduction. *Ecology letters*, 18(1):74–84, 2015.
- Guy Latouche, Marie-Ange Remiche, Peter Taylor, et al. Transient markov arrival processes. *The Annals of Applied Probability*, 13(2):628–640, 2003.
- Jostein Starrfelt and Hanna Kokko. Bet-hedging—a triple trade-off between means, variances and correlations. *Biological Reviews*, 87(3):742–755, 2012.
- Hiroshi Toyoizumi and Jeremy Field. Dynamics of social queues. *Journal of Theoretical Biology*, 346(0):16 – 22, 2014. ISSN 0022-5193. doi: <http://dx.doi.org/10.1016/j.jtbi.2013.12.017>. URL <http://www.sciencedirect.com/science/article/pii/S0022519313005687>.

Correlations of Interference in Non-Poisson Networks under Spatially Correlated Shadowing

Tatsuaki Kimura

Department of Information and Communications Technology,
Graduate School of Engineering,
Osaka University

Abstract In this paper, we theoretically analyze temporal correlations of interference under a spatially correlated shadowing environment in non-Poisson networks. We utilize well-accepted Gudmundson (1991)'s model for the spatial correlation of shadowing and model the transmitter locations by *repulsive* point processes and *attractive* one. For the repulsive one, we consider determinantal point processes and for the attractive one, doubly Poisson cluster processes. In addition, we consider i.i.d. mobility models for each transmitter. By using a useful mathematical tool, Watson's lemma, we derive simple asymptotic formulas of the temporal correlation coefficient of interference when the variance of the shadowing is large. The obtained results show the readable relationship between the correlation of shadowing, that of transmitter locations, mobility effect, and the correlation of interference.

1. Introduction

In general, shadowing effect in wireless transmission channels is spatially correlated on a scale from 50 to 200 m [1] due to the effects of blockage and reflection. A well-accepted model for the correlation of shadowing is Gudmundson's model [2], in which the shadowing is modeled as an autoregressive log-Gaussian process with an exponentially decaying autocorrelation function subject to the moving distance of the transmitting node. Due to the spatial and temporal aspects of wireless systems, such spatial correlation may affect the performance of wireless communications. For example, if the channel quality is spatially correlated, outages at multiple antennas are spatially correlated, which may lead to frequent simultaneous outages. However, few studies theoretically analyzed the impact of the spatially correlated shadowing on the performance of wireless communication systems due to its mathematical intractability. Most of the previous works assumed i.i.d. fading (only Rayleigh fading is often assumed). In addition, the impacts of the mobility of transmitters in the correlated shadowing environment has not been sufficiently analyzed.

To characterize the spatial and temporal performance of wireless systems, correlations of interference or outages have been studied in this decade. Indeed, interference is one of key factors of the performance of various wireless communication systems because it directly relates to the communication quality. In addition, interference can be spatially and temporally correlated due to the locations of interferers, transmission patterns, and fading effect including the spatially correlated shadowing. Furthermore, such correlations of interference may degrade the performance of wireless communications in various aspects, e.g., end-to-end throughput, retransmissions [3], cooperative relaying [4], broadcast communications, multiple antennas [5], and handovers [6]. For example, quick retransmission after a transmission failure is likely to fail again due to the temporal correlation of interference. Thus, understanding the correlation of interference is important in the design and control of various wireless systems.

In this paper, we study the temporal correlation of interference in mobile ad-hoc networks (MANETs) under a spatially correlated shadowing environment. We model the locations of transmitters (i.e., interferers) by two types of non-Poisson point processes: *repulsive* one and *attractive* one. More precisely, we consider determinantal point processes and doubly Poisson cluster processes for repulsive and attractive point processes, respectively. The spatially correlated shadowing is modeled by widely-accepted Gudmundson's model [2], and we consider i.i.d. mobility models for the mobility of transmitters. In our previous work, we study the temporal correlation of interference under spatially correlated shadowing in Poisson networks [7]. Thus, this paper is an extension of the work toward non-Poisson network models. The Poisson model assumes independence among the locations of transmitters, however, they may have positive/negative correlations in the real world due to many reasons, such as geographical limitations and human

activity. Therefore, we reveal the impacts of such correlations of transmitter locations and shadowing on the performance of wireless communications in this work. However, the temporal correlation coefficient does not have a closed-form expression due to these correlations. Thus, we derive simple asymptotic expansions of the correlation coefficients when a *variance* of the log-normal shadowing increases by enhancing the approach in [7]. The obtained results show the readable relationship between the correlations of the shadowing, transmitter locations, mobility and the correlation of interference.

In what follows, we briefly summarize related works in the area of theoretical analysis of spatial or temporal correlations in wireless communications. Due to their importance, many researchers considered spatial and temporal correlation of interference in various scenarios in this decade [6, 7, 8, 9, 10, 11, 12, 13, 14]. Those in Poisson ad-hoc networks were first studied by Ganti and Haenggi [8], and they showed that interference can be correlated due to node locations even with ALOHA. Schilcher et al. [9] extended this work by considering the durations of traffic and block fading. The impact of the mobility of transmitters on the correlation of interference was studied in [10]; and it was shown that the correlation of interference becomes lower in a higher mobility environment. As extensions of a model of node locations, Wen et al. considered K -tier heterogeneous networks [11] and also studied the cases when node locations are distributed with a Matérn cluster process or a second-order cluster process that has a repulsive property. Krishnan and Dhillon [6] studied the spatial and temporal correlation of interference and the joint coverage probability in a cellular network with a closest base station association policy. Furthermore, Koufos and Dettmann [15] studied the impact of the correlation of the node mobility with a random waypoint mobility model in a 1-D finite lattice. However, these studies assumed i.i.d. fading for each transmission channel (often Rayleigh fading is assumed) due to mathematical tractability and ignored the spatial correlation of shadowing. Most recently, Koufos et al. [13, 14] analyzed the effect of blockage on the temporal correlation of interference. In their latest work [14], they modeled the locations of nodes and obstacles by 1-D PPPs and consider the penetration loss of obstacles. However, they did not analyze the relationship between the correlation distance of shadowing and correlation of interference in a general framework. In addition, the obtained results are complicated, and so the impacts of various system parameters were not clearly shown even though their modeling assumption was quite simple.

There are also several studies that aimed to efficiently model or approximate the mathematically intractable shadowing effect. Błaszczyszyn and Karray [16] studied the mean interference in cellular networks when the variance of the shadowing increases. Similarly, Błaszczyszyn et al. [17] considered the path-loss process. Heath et al. [18] approximated the log-normal shadowing by gamma distributions. Whereas the above studies assumed i.i.d. fading for each transmission channel, Renzo et al. [19] considered a model where all channels are equicorrelated. Baccelli and Zhang [20] recently proposed a correlated shadowing model in which correlated log-normal shadowing is approximated by a random variable depending on the number of buildings that a transmission channel penetrates. Recently, this idea was applied to a 3-D network model by Lee et al. [21]. Contrary to the above studies, we aim to derive a simple and readable relationship between the spatial correlation of shadowing and temporal correlation of interference on the basis of Gudmundson's shadowing model. In addition, we also investigate the impacts of node mobility and correlation of node locations in this paper.

The reminder of this paper is organized as follows. Section 2 explains our model considered in this paper. Section 3 presents our main results, the temporal correlation of interference. Finally, we conclude this paper in Section 4.

2. Model Description

2.1. Network modeling

We consider 2D-space $\mathbb{R}^2$, in which the target receiver is located at the origin o , and assume that time is slotted as $t \in \mathbb{Z}_+ \triangleq \{0, 1, 2, \dots\}$. In addition, potential transmitters (interferers)¹ are randomly distributed on the space at the initial time slot $t = 0$ according to a certain point process and randomly move in each time slot according to an i.i.d. mobility model. Let $x_i(t) \in \mathbb{R}^2$ denote the location of the i -th transmitter ($i \in \mathbb{Z}_+$)

¹In this paper, we only consider the transmitters of the interference channel and assume that a certain transmitter of the desired channel is located independent of the interferers.

and let $\Phi(t) = \{x_i(t); i \in \mathbb{Z}_+\}$ denote the locations of transmitters at the time slot t . In this paper, we consider two types of non-Poisson point processes for the initial transmitters locations (i.e., $\Phi(0)$): *repulsive* point processes and *attractive* one. More precisely, we consider a *determinantal point process (DPP)* for the repulsive one and *doubly Poisson cluster process (DPCP)* for the attractive one. In what follows, we give the detailed definitions of these point processes. To simplify the notation, we write $\Phi(0) \equiv \Phi$, $x_i(0) \equiv x_i$ when considering only the initial time slot $t = 0$.

2.1.1. Determinantal point process (DPP)

We first define the n -th moment density function $\rho^{(n)}(x_1, \dots, x_n)$ of Φ such that for $B_1, \dots, B_n \subset \mathbb{R}^2$,

$$\int_{B_1} \dots \int_{B_n} \rho^{(n)}(x_1, \dots, x_n) dx_1 \dots dx_n = \mathbb{E} \left[\sum_{x_1, \dots, x_n \in \Phi}^{\neq} \mathbb{1}(x_1 \in B_1) \dots \mathbb{1}(x_n \in B_n) \right], \quad (1)$$

where $\mathbb{1}(\cdot)$ denotes the indicator function and $\neq$ indicates the sum is taken over distinct points on Φ . The DPPs are characterized by their moment density function:

$$\rho^{(n)}(x_1, \dots, x_n) = \det[K](x_1, \dots, x_n),$$

where K is a kernel function. In addition, for stationary DPPs, $\rho^{(2)}(x, y)$ satisfies

$$\rho^{(2)}(x, y) = \rho_0(\|x - y\|), \quad x, y \in \mathbb{R}^2,$$

where $\|\cdot\|$ denotes the Euclidean distance. A well-known example of DPPs is a β -Ginibre point process (β -GPP). Formally, if K is the Gaussian kernel such that

$$K_{\beta,c}(x, y) = \frac{c}{\pi} e^{-c(|x|^2 + |y|^2)/(2\beta)} e^{(c/\beta)x\bar{y}}, \quad x, y \in \mathbb{C},$$

Φ is called as the β -GPP with intensity $\lambda = c\pi^{-1}$ ($c > 0$). Its second moment density function is equal to

$$\rho^{(2)}(x_1, x_2) = \left(\frac{c}{\pi}\right)^2 (1 - e^{-\frac{c}{\beta}\|x_1 - x_2\|^2}), \quad x_1, x_2 \in \mathbb{R}^2. \quad (2)$$

2.1.2. Doubly Poisson cluster process (DPCP)

In general, a cluster process is generated from a parent point process and related daughter point processes (i.e., a point process in a cluster around each parent). In the DPCP, the parent point process is a homogeneous PPP with intensity λ_p , and the daughter points in each cluster are distributed according to an i.i.d. finite PPP with intensity $\lambda_c(y)$ around the parent point with mean $\bar{c}$. Clearly, the intensity of the DPCP is $\lambda = \lambda_p \bar{c}$.

An example of DPCPs is a Thomas cluster process (TCP), in which each cluster has Gaussian intensity function

$$\lambda_c(x) = \frac{\bar{c}}{2\pi\sigma_c^2} \exp\left(-\frac{\|x\|^2}{2\sigma_c^2}\right). \quad (3)$$

2.2. Mobility modeling

In this paper, we consider the following i.i.d. mobility model. The position of the transmitter i at the time slot $(t + 1)$ is determined as $x_i(t + 1) = x_i(t) + v_i$ and v_i is independently distributed with a probability density function $\psi(v)$ ($v \in \mathbb{R}^2$). Thus, the moving distance of the transmitter i in τ time-slots, i.e., $v_i(\tau) \triangleq x_i(\tau) - x_i(0)$, has the p.d.f. $\psi_\tau(v) \triangleq \psi^{\tau*}(v)$, which represents the τ -th convolution of $\psi(v)$. In addition, we assume that $\bar{\psi}(v)$ depends only on the moving distance, i.e., $\psi(v) \equiv \bar{\psi}(\|v\|)$ and $\psi_\tau(v) \equiv \bar{\psi}_\tau(\|v\|)$. We choose the following two models as examples of $\psi(v)$: (i) *constrained i.i.d. mobility model*; and (ii) *discrete-time Brownian motion model*. The same models are considered in [10].

(i) *Constrained i.i.d. mobility (CIM) model*: In the CIM model, the location $x_i(t + 1) \in \Phi(t + 1)$ of the transmitter i at the time slot $(t + 1)$ is determined independent of $x_i(t)$ such that

$$x_i(t + 1) := x_i(0) + v_i(t), \quad t \in \mathbb{Z}_+.$$

Here, $v_i(t)$'s are uniformly distributed in $B(o, V_{\max})$, a ball centered at o with radius $V_{\max}$. Clearly,

$$\psi_\tau(v) = \begin{cases} \frac{1}{\pi V_{\max}^2}, & v \in B(o, V_{\max}), \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

(ii) *Discrete-time Brownian motion model (BM)*: In the BM model, the location of the transmitter i at the time slot $(t + 1)$ is determined by

$$x_i(t + 1) := x_i(t) + v_i(t), \quad t \in \mathbb{Z}_+, \quad (5)$$

with $v_i(t) \sim \mathcal{N}(\mathbf{0}, \sigma_V^2 \mathbf{I})$, and thus $\psi_\tau(v) = \mathcal{N}(\mathbf{0}, \tau \sigma_V^2 \mathbf{I})$.

2.3. Channel modeling

We next explain our channel model. For simplicity, we assume that all transmitters have the unit transmission power. The distance based path loss model is assumed as

$$\ell(r) = (\varepsilon_0 + |r|)^\alpha, \quad r \in \mathbb{R},$$

where $\alpha > 2$ is a path loss exponent and ε_0 denotes a parameter for avoiding singularity at $r = 0$. Let h denote the shadowing variable. According to a widely-accepted assumption for the shadowing effect [1], h is assumed to be log-normally distributed,

$$h = \exp(-\sigma_{\text{dB}}^2/2 + \sigma_{\text{dB}} Z),$$

where $Z \sim \mathcal{N}(0, 1)$. Note that $\mathbb{E}[h] = 1$ and $\mathbb{E}[h^2] = e^{\sigma_{\text{dB}}^2}$. In addition, we model the spatial correlation of shadowing on the basis of Gudmundson [2]'s model as follows. If we consider two transmitters at x_i and x_j ($i \neq j$) in $\mathbb{R}^2$ and their transmissions to o , the corresponding shadowing variables h_i and h_j have the following correlation coefficient in the logarithmic sense depending on the distance between the transmitters:

$$\rho_{\text{dB}} \triangleq \frac{\mathbb{E}[\ln h_i \ln h_j]}{\sigma_{\text{dB}}^2} = e^{-\frac{\|x_i - x_j\|}{d_{\text{cor}}} \ln 2}. \quad (6)$$

Here, d_{cor} is known as the correlation distance and a typical value of d_{cor} in urban environments is [50, 200] m (see e.g., [2]). For simplicity, let $d_0 \equiv d_{\text{cor}} / \ln 2$. Therefore, (6) can be rewritten as

$$\mathbb{E}[h_i h_j] = \exp(\sigma_{\text{dB}}^2 \rho_{\text{dB}}) = \exp\left(\sigma_{\text{dB}}^2 e^{-\frac{\|x_i - x_j\|}{d_0}}\right). \quad (7)$$

In this paper, we only consider the effect of the shadowing and do not take into account the multi-path fading, such as Rayleigh fading. In addition, the shadowing is assumed to be time-invariant, i.e., shadowing variables do not depend on time but geographical locations of transmitters.

By definition, the received power from the transmitter i at the time slot t can be represented as $h_i(t)/\ell(x_i(t))$. Thus, the total interference power received at o is

$$I(t) = \sum_{x_i(t) \in \Phi(t)} \frac{h_i(t) \mathcal{S}_i(t)}{\ell(x_i(t))}, \quad (8)$$

where $\mathcal{S}_i(t)$ denotes an i.i.d Bernoulli random variable with mean q and equals to 1 when the transmitter i is transmitting radio waves at the beginning of the time slot t , otherwise 0. In other words, we assume that each transmitter transmits radio waves with probability q in each transmission time-slot. In addition, we assume that the transmission period is fixed and smaller than the mobility time-slot and thus the transmission of each transmitter at the time slot t does not continue until the next time slot $(t + 1)$. Therefore, $\mathcal{S}_i(t)$ and $\mathcal{S}_i(t + 1)$ are independent for all i .

3. Main Results

In this section, we present our main results, the temporal correlation of interference. To do this, we first derive the first and second moments of interference, which will be used for the derivation of the temporal correlation coefficient of interference. Due to the correlations of shadowing effect and transmitter locations, the second moment and temporal correlation coefficient do not have explicit forms. Therefore, by applying Watson's lemma (see e.g., [24] and Appendix A), we derive their simple asymptotic expansions for sufficiently large σ_{dB} .

In what follows, we use the order notation as below. Let $f(x)$ and $g(x)$ denote functions on $\mathbb{R}$. We write

$$f(x) = O(g(x)), \quad \text{as } x \rightarrow \infty,$$

if there exists some $M > 0$ such that

$$f(x) = O(g(x)), \quad \text{for } x > M.$$

3.1. First and second moments of interference

The mean interference can be easily obtained from Campbell-Mecke's theorem (see e.g., [23]):

$$\mathbb{E}[I] = \mathbb{E} \left[\sum_{x_i \in \Phi} \frac{h_i \mathcal{S}_i}{\ell(x_i)} \right] = \int_{\mathbb{R}^2} \frac{\lambda q \mathbb{E}[h]}{\varepsilon_0 + \|x\|^\alpha} dx = \frac{2\pi^2 \lambda q}{\varepsilon_0^{1-2/\alpha} \alpha \sin(2\pi/\alpha)}. \quad (9)$$

We next consider the second moment of the interference, $\mathbb{E}[I^2]$. By definition, this can be rewritten as

$$\mathbb{E}[I^2] = \mathbb{E} \left[\sum_{x_i, x_j \in \Phi}^{\neq} \frac{h_i h_j \mathcal{S}_i \mathcal{S}_j}{\ell(x_i) \ell(x_j)} \right] + \mathbb{E} \left[\sum_{x_i \in \Phi} \frac{h_i^2 \mathcal{S}_i}{(\ell(x_i))^2} \right]. \quad (10)$$

The second term in the above can be easily computed by applying Campbell-Mecke's theorem (e.g., [23]) as follows.

$$\mathbb{E} \left[\sum_{x_i \in \Phi} \frac{h_i^2 \mathcal{S}_i}{(\ell(x_i))^2} \right] = \int_{\mathbb{R}^2} \frac{\lambda q \mathbb{E}[h^2]}{(\varepsilon_0 + \|x\|^\alpha)^2} dx = \lambda q e^{\sigma_{\text{dB}}^2} \gamma, \quad (11)$$

where

$$\gamma = \frac{2\pi^2(\alpha - 2)}{\varepsilon_0^{2-2/\alpha} \alpha^2 \sin(2\pi/\alpha)}. \quad (12)$$

On the other hand, the first term in (10) can be numerically computed, but it does not have a closed-form due to the spatial correlation of shadowing and that of transmitter locations. Lemma 1 below shows simple asymptotic expansions of the second moment of interference when the variance of log-normal shadowing (i.e., σ_{dB}) is large.

Lemma 1 *The second moment of interference has the following asymptotic expansions as $\sigma_{\text{dB}}^2 \rightarrow \infty$: (i) If Φ is the DPP with the second moment density function $\rho_0(r)$,*

$$\mathbb{E}[I^2] = q e^{\sigma_{\text{dB}}^2} \gamma \left[\lambda + \frac{4\pi d_0^3 q \rho'_0(0)}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad (13)$$

and (ii) if Φ is DPCP with the cluster intensity function $\lambda_c(r)$

$$\mathbb{E}[I^2] = q \lambda e^{\sigma_{\text{dB}}^2} \gamma \left[\lambda + \frac{D_{\text{DPCP}}^{(4)}}{\sigma_{\text{dB}}^4} + \frac{D_{\text{DPCP}}^{(6)}}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad (14)$$

where

$$D_{\text{DPCP}}^{(4)} = 2\pi d_0^2 q \left(\lambda + \frac{1}{c} \lambda_c^{(2*)}(0) \right), \quad D_{\text{DPCP}}^{(6)} = 2\pi d_0^2 q \left(3\lambda + 3\frac{1}{c} \lambda_c^{(2*)}(0) + \frac{2d_0}{c} \lambda_c * \lambda'_c(0) \right). \quad (15)$$

Proof. The proof of this lemma is given in Appendix A. $\square$

Corollary 1 *The second moment of interference has the following asymptotic expansions as $\sigma_{\text{dB}} \rightarrow \infty$: (i) If Φ is the β -GPP,*

$$\mathbb{E}[I^2] = \lambda q e^{\sigma_{\text{dB}}} \gamma \left[1 + \frac{12\pi d_0^4 q c}{\beta} \frac{1}{\sigma_{\text{dB}}^8} + O\left(\frac{1}{\sigma_{\text{dB}}^{10}}\right) \right],$$

and (ii) if Φ is the TCP,

$$\mathbb{E}[I^2] = \lambda q e^{\sigma_{\text{dB}}} \gamma \left[1 + 2\pi d_0^2 \lambda q \left(1 + \frac{1}{4\pi \lambda_p \sigma_c^2} \right) \left\{ \frac{1}{\sigma_{\text{dB}}^4} + \frac{3}{\sigma_{\text{dB}}^6} \right\} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right].$$

Proof. The proof of this corollary is given in Appendix C. $\square$

3.2. Temporal correlation coefficient

In this section, we derive the temporal correlation of interference. To do this, we first give the definitions of the *temporal covariance* and *temporal correlation coefficient* of interference. The temporal covariance of interference for time interval τ slots ($\tau \in \mathbb{Z}_+$) is given by

$$\text{Cov}[I(0), I(\tau)] = \mathbb{E}[I(0)I(\tau)] - (\mathbb{E}[I])^2. \quad (16)$$

Similarly, we define the temporal correlation coefficient of interference as follows.

$$\rho_{\text{tm}, \tau} \triangleq \frac{\text{Cov}[I(0), I(\tau)]}{\sqrt{\text{Var}[I(0)]\text{Var}[I(\tau)]}} = \frac{\mathbb{E}[I(0)I(\tau)] - (\mathbb{E}[I])^2}{\mathbb{E}[I^2] - (\mathbb{E}[I])^2}.$$

The following results are asymptotic expressions of the temporal covariance.

Lemma 2 *Suppose $\bar{\psi}'_\tau(0)$ exists. In addition, for any function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, we write*

$$\mathbb{E}_{\psi_\tau}[f(v)] = \int_{\mathbb{R}^2} f(v) \psi_\tau(v) dv.$$

The temporal covariance of interference for time interval τ has the following asymptotic expansion as $\sigma_{\text{dB}}^2 \rightarrow \infty$: (i) If Φ is the DPP,

$$\begin{aligned} \text{Cov}[I(0), I(\tau)] &= 2\pi q^2 d_0^2 e^{\sigma_{\text{dB}}} \gamma \left[\frac{\lambda \bar{\psi}_\tau(0) + \mathbb{E}_{\psi_\tau}[\rho_0(\|v\|)]}{\sigma_{\text{dB}}^4} \right. \\ &\quad \left. + \frac{3\lambda \bar{\psi}_\tau(0) + 2\lambda d_0 \bar{\psi}'_\tau(0) + 3\mathbb{E}_{\psi_\tau}[\rho_0(\|v\|)]}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \end{aligned} \quad (17)$$

and (ii) if Φ is the DPCP,

$$\text{Cov}[I(0), I(\tau)] = q \lambda e^{\sigma_{\text{dB}}} \gamma \left[\frac{C_{\text{DPCP}}^{(4)}}{\sigma_{\text{dB}}^4} + \frac{C_{\text{DPCP}}^{(6)}}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad (18)$$

where

$$C_{\text{DPCP}}^{(4)} = 2\pi d_0^2 q \left(\bar{\psi}_\tau(0) + \lambda + \frac{1}{c} \mathbb{E}_{\psi_\tau}[\lambda_c^{(2*)}(\|v\|)] \right), \quad (19)$$

$$C_{\text{DPCP}}^{(6)} = 2\pi d_0^2 q \left(3(\bar{\psi}_\tau(0) + \lambda + \frac{1}{c} \mathbb{E}_{\psi_\tau}[\lambda_c^{(2*)}(v)]) + 2d_0 \bar{\psi}'_\tau(0) \right). \quad (20)$$

Proof. The proof of this lemma is omitted due to the space limitations. $\square$

We can easily confirm that $\bar{\psi}'_\tau(0) = 0$ in our mobility models. In addition, $\bar{\psi}_\tau(0) = 1/(\pi V_{\text{max}}^2)$ in the CIM model and $\bar{\psi}_\tau(0) = 1/(2\pi\tau\sigma_V^2)$ in the BM model (see Section 2.2). The terms $\mathbb{E}_{\psi_\tau}[\rho_0(\|v\|)]$ and $\mathbb{E}_{\psi_\tau}[\lambda_0^{(2*)}(\|v\|)]$ also can be calculated by using (2) and (3). As a result, we obtain the following corollary.

Corollary 2 *The temporal covariance has the following asymptotic expansions as $\sigma_{\text{dB}} \rightarrow \infty$:*

$$\text{Cov}[I(0), I(\tau)] = 2\pi\lambda q^2 d_0^2 e^{\sigma_{\text{dB}}^2} \gamma \left[\frac{C_V + \lambda C_\rho}{\sigma_{\text{dB}}^4} + \frac{3C_V + 3\lambda C_\rho}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right],$$

where (i) in the β -GPP model with the CIM mobility model,

$$C_V = \frac{1}{\pi V_{\text{max}}^2}, \quad C_\rho = \frac{\beta}{c V_{\text{max}}^2} \left(\frac{c}{\beta} + e^{-\frac{c}{\beta} V_{\text{max}}^2} - 1 \right),$$

(ii) in the β -GPP model with the BRM mobility model,

$$C_V = \frac{1}{2\pi\tau\sigma_V^2}, \quad C_\rho = \frac{2c\tau\sigma_V^2}{\beta + 2c\tau\sigma_V^2},$$

(iii) in the TCP model with the CIM mobility model,

$$C_V = \frac{1}{\pi V_{\text{max}}^2}, \quad C_\rho = 1 + \frac{1}{\lambda_p \pi V_{\text{max}}^2} \left(1 - e^{-\frac{V_{\text{max}}^2}{4\sigma_c^2}} \right),$$

and (iv) in the TCP model with the BRM mobility model,

$$C_V = \frac{1}{2\pi\tau\sigma_V^2}, \quad C_\rho = 1 + \frac{1}{2\pi\lambda_p(2\sigma_c^2 + \tau\sigma_V^2)}.$$

Finally, we provide asymptotic expressions of the temporal correlation coefficient of interference.

Theorem 1 *Suppose $\bar{\psi}'_\tau(0)$ exists. The temporal correlation coefficient of interference of time interval τ is asymptotically equivalent to when $\sigma_{\text{dB}} \rightarrow \infty$: (i) If Φ is the DPP,*

$$\begin{aligned} \rho_{\text{tm},\tau} &= 2\pi q d_0^2 \left[1 + \frac{1}{\lambda} \frac{4\pi d_0^3 q \rho'_0(0)}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right]^{-1} \\ &\times \left[\frac{\bar{\psi}_\tau(0) + \frac{1}{\lambda} \mathbb{E}_{\psi_\tau}[\rho_0(\|v\|)]}{\sigma_{\text{dB}}^4} + \frac{3\bar{\psi}_\tau(0) + 2d_0\bar{\psi}'_\tau(0) + \frac{3}{\lambda} \mathbb{E}_{\psi_\tau}[\rho_0(\|v\|)]}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \end{aligned} \quad (21)$$

and (ii) if Φ is the DPCP,

$$\rho_{\text{tm},\tau} = \left[1 + \frac{D_{\text{DPCP}}^{(4)}}{\sigma_{\text{dB}}^4} + \frac{D_{\text{DPCP}}^{(6)}}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right]^{-1} \left[\frac{C_{\text{DPCP}}^{(4)}}{\sigma_{\text{dB}}^4} + \frac{C_{\text{DPCP}}^{(6)}}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad (22)$$

where $D_{\text{DPCP}}^{(4)}$, $D_{\text{DPCP}}^{(6)}$, $C_{\text{DPCP}}^{(4)}$, and $C_{\text{DPCP}}^{(6)}$ are given in (15), (19), and (20), respectively.

Corollary 3 *The temporal correlation coefficient is asymptotically equivalent to as $\sigma_{\text{dB}} \rightarrow \infty$: (i) If Φ is the β -GPP,*

$$\rho_{\text{tm},\tau} = 2\pi q d_0^2 \left(1 + \frac{12\pi d_0^4 q c}{\beta} \frac{1}{\sigma_{\text{dB}}^8} + O\left(\frac{1}{\sigma_{\text{dB}}^{10}}\right) \right)^{-1} \left[\frac{C_V + \lambda C_\rho}{\sigma_{\text{dB}}^4} + \frac{3C_V + 3\lambda C_\rho}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad (23)$$

and (ii) if Φ is the TCP,

$$\begin{aligned} \rho_{\text{tm},\tau} &= 2\pi q d_0^2 \left(1 + 2\pi d_0^2 \lambda q \left(1 + \frac{1}{4\pi\lambda_p\sigma_c^2} \right) \left\{ \frac{1}{\sigma_{\text{dB}}^4} + \frac{3}{\sigma_{\text{dB}}^6} \right\} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right)^{-1} \\ &\times \left[\frac{C_V + \lambda C_\rho}{\sigma_{\text{dB}}^4} + \frac{3C_V + 3\lambda C_\rho}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \end{aligned} \quad (24)$$

where C_V and C_ρ are defined in Corollary 2.

4. Conclusion

In this paper, we study the temporal correlations of interference in non-Poisson networks under spatially correlated shadowing. By assuming Gudmundson's correlated shadowing model [2] and modeling transmitter locations by determinantal point processes (DPPs) and doubly Poisson cluster processes (DPCPs), we derive the temporal correlation coefficient for each model. Since the exact expressions of the temporal correlation coefficients do not have closed forms, we derive simple asymptotic expansions of them when the variance of the log-normal shadowing is large. The obtained results show the readable relationship between system parameters. Since we only focus on the correlations of interference in this work, spatio-temporal analysis of more practical performance metrics, such as correlations of outage probabilities under spatially correlated shadowing, is our future work.

A. Watson's Lemma

In this appendix, we explain Watson's lemma, which is a useful mathematical tool in an asymptotic analysis of the following form of exponential integrals:

$$\tilde{F}(\sigma) \triangleq \int_a^{a+\delta} e^{\sigma R(t)} g(t) dt = \int_0^\delta e^{\sigma R(a+\tau)} g(a+\tau) d\tau, \quad (25)$$

where $\delta > 0$, $R(t)$ has its maximum at $t = a$, and $g(t)$ is a nonnegative function. Moreover, let $\tilde{R}(t) := R(t) - R(a)$. Watson's lemma gives an asymptotic expression of $\tilde{F}(\sigma)$ when σ is large.

Proposition 1 (Miller [24] (Section 3.3)) *Suppose that $R(t)$ and $g(t)$ have an infinite number of continuous derivatives for $a \leq t < a + \delta$ and $R'(a) < 0$. We then have the following asymptotic expansion:*

$$\tilde{F}(\sigma) = \sum_{n=0}^{\infty} \frac{\psi^{(n)}(0)}{\sigma^{n+1}}, \quad \text{as } \sigma \rightarrow \infty \text{ with } \sigma > 0, \quad (26)$$

where

$$\psi(s) = g(a + \tau(s))\tau'(s), \quad \tilde{R}(a + \tau) = -s, \quad (27)$$

We can immediately obtain the following corollary from Proposition 1.

Corollary 4

$$\tilde{F}(\sigma) = e^{\sigma R(a)} \left[\frac{\psi(0)}{\sigma} + \frac{\psi'(0)}{\sigma^2} + \frac{\psi''(0)}{\sigma^3} + \frac{\psi'''(0)}{\sigma^4} + O\left(\frac{1}{\sigma^5}\right) \right], \quad \text{as } \sigma \rightarrow \infty,$$

where

$$\begin{aligned} \psi(0) &= -\frac{g(a)}{R'(a)}, & \psi'(0) &= \frac{1}{(R'(a))^3} (g'(a)R'(a) - g(a)R''(a)), \\ \psi''(0) &= -\frac{1}{(R'(a))^5} [g''(a)(R'(a))^2 - 3g'(a)R'(a)R''(a) - g(a)\{R'(a)R'''(a) - 3(R''(a))^2\}], \\ \psi'''(0) &= \frac{1}{(R'(a))^7} [g'''(a)(R'(a))^3 - 6g''(a)(R'(a))^2R''(a) + g'(a)R'(a)\{15(R''(a))^2 - 4R'(a)R'''(a)\} \\ &\quad - g(a)\{(R''(a))^2R'''(a) - 10R'(a)R''(a)R'''(a) + 15(R''(a))^3\}]. \end{aligned} \quad (28)$$

Furthermore, by applying the above results, we obtain Proposition 2 below, which is used in the proof of our main results.

Proposition 2 (Proposition 4.2 in [25]) *Let $f(\cdot)$ denote an arbitrary function on $\mathbb{R}$ such that $f'(0)$ exists. In addition, let*

$$\Psi(f) = \iint_{(\mathbb{R}^2)^2} \frac{\exp\left(\sigma^2 e^{\frac{-\|y\|}{d_0}}\right) f(\|y\|)}{\ell(\|x\|)\ell(\|x+y\|)} dy dx.$$

We then have

$$\Psi(f) = 2\pi d_0^2 e^{\sigma^2} \gamma \left[\frac{f(0)}{\sigma^4} + \frac{3f(0) + 2d_0 f'(0)}{\sigma^6} + O\left(\frac{1}{\sigma^8}\right) \right], \quad \text{as } \sigma \rightarrow \infty,$$

where γ is given in (12).

B. Proof of Lemma 1

We start with the case of the DPP. It follows from (1) and Campbell's theorem (see e.g., [23]) that

$$\begin{aligned} \Gamma &\triangleq \mathbb{E} \left[\sum_{x_i, x_j \in \Phi}^{\neq} \frac{h_i h_j \mathcal{S}_i \mathcal{S}_j}{\ell(\|x_i\|) \ell(\|x_j\|)} \right] = q^2 \iint_{(\mathbb{R}^2)^2} \frac{\mathbb{E}_{ij}[h_i h_j]}{\ell(\|x_i\|) \ell(\|x_j\|)} \rho^{(2)}(x_i, x_j) dx_i dx_j \\ &\stackrel{(a)}{=} q^2 \iint_{(\mathbb{R}^2)^2} \frac{e^{\sigma_{\text{dB}}^2} e^{-\frac{\|y\|}{d_0}}}{\ell(\|x\|) \ell(\|x+y\|)} \rho_0(\|y\|) dx dy, \end{aligned} \quad (29)$$

where we use (7) in (a). Note here that $\rho_0(0) = 0$ in the DPP. Therefore, by applying Proposition 2 to the above, we obtain

$$\Gamma = 2\pi q^2 d_0^2 e^{\sigma_{\text{dB}}^2} \gamma \left[\frac{2d_0 \rho'_0(0)}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad \text{as } \sigma_{\text{dB}} \rightarrow \infty. \quad (30)$$

Combining the above and (11) with (10) results in (13).

We next consider the case of the DPCP. The second moment density function of the DPCP with cluster intensity function $\lambda_c(r)$ is given by (see e.g., [23])

$$\rho_0^{(2)}(r) = \lambda^2 + \lambda_p \lambda_c^{(2*)}(r), \quad r \in \mathbb{R}_+,$$

where $f^{(2*)}$ represents the two times self-convolution of a function f . Similar to the derivation of (30), applying Campbell's theorem (see e.g., [23]) to (1) yields

$$\Gamma = q^2 \iint_{(\mathbb{R}^2)^2} \frac{e^{\sigma_{\text{dB}}^2} e^{-\frac{\|y\|}{d_0}}}{\ell(\|x\|) \ell(\|x+y\|)} \left(\lambda^2 + \lambda_p \lambda_c^{(2*)}(\|y\|) \right) dx dy.$$

It follows from Proposition 2 that

$$\begin{aligned} \Gamma &= 2\pi q^2 d_0^2 e^{\sigma_{\text{dB}}^2} \gamma \left[\left(\lambda^2 + \lambda_p \lambda_c^{(2*)}(0) \right) \frac{1}{\sigma_{\text{dB}}^4} \right. \\ &\quad \left. + \left(3\lambda^2 + 3\lambda_p \lambda_c^{(2*)}(0) + 2d_0 \lambda_p \lambda_c * \lambda'_c(0) \right) \frac{1}{\sigma_{\text{dB}}^6} + O\left(\frac{1}{\sigma_{\text{dB}}^8}\right) \right], \quad \text{as } \sigma_{\text{dB}} \rightarrow \infty. \end{aligned} \quad (31)$$

Finally, substituting the above and (11) into (10) leads to (14).

C. Proof of Corollary 2

The statement (ii) can be easily proved by applying Lemma 1, however, we cannot use Lemma 1 in the case of the β -GPP because $\rho_0(0) = \rho'_0(0) = 0$. Thus, we prove this case by modifying the proof of Lemma 1.

By considering a planar coordination, i.e., $x := (r \cos \theta, r \sin \theta)$, $y := (s \cos \phi, s \sin \phi)$, Γ in (29) can be rewritten as

$$\Gamma = q^2 \int_0^{2\pi} \int_{\mathbb{R}_+} \int_0^{2\pi} \int_{\mathbb{R}_+} \frac{s r e^{\sigma_{\text{dB}}^2} e^{-\frac{s}{d_0}} \rho_0(s) ds d\phi dr d\theta}{\ell(r) \ell(\sqrt{r^2 + s^2 + 2sr \cos(\phi - \theta)})}. \quad (32)$$

Recall that $\rho_0(0) = \rho'_0(0) = 0$ in the β -GPP. Since the function $e^{-\frac{s}{d_0}}$ has its maximum at $s = 0$, it follows from Corollary 4 that as $\sigma_{\text{dB}} \rightarrow \infty$

$$\int_{\mathbb{R}_+} \frac{s r e^{\sigma_{\text{dB}}^2} e^{-\frac{s}{d_0}} \rho_0(s)}{\varepsilon_0 + (r^2 + s^2 + 2sr \cos(\phi - \theta))^{\frac{\alpha}{2}}} ds = d_0^4 e^{\sigma_{\text{dB}}^2} r \left[\frac{3\rho_0''(0)}{\varepsilon_0 + r^\alpha} \frac{1}{\sigma_{\text{dB}}^8} + O\left(\frac{1}{(\varepsilon_0 + r^\alpha) \sigma_{\text{dB}}^{10}}\right) \right],$$

which leads to

$$\begin{aligned}\Gamma &= q^2 d_0^4 e^{\sigma_{\text{dB}}^2} \iint d\phi d\theta \int_{\mathbb{R}_+} r \left[\frac{3\rho_0''(0)}{(\varepsilon_0 + r^\alpha)^2} \frac{1}{\sigma_{\text{dB}}^8} + O\left(\frac{1}{(\varepsilon_0 + r^\alpha)^2 \sigma_{\text{dB}}^{10}}\right) \right] \\ &= 2\pi q^2 d_0^4 e^{\sigma_{\text{dB}}^2} \gamma \left[\frac{3\rho_0''(0)}{\sigma_{\text{dB}}^8} + O\left(\frac{1}{\sigma_{\text{dB}}^{10}}\right) \right].\end{aligned}$$

In addition, $\rho_0''(0) = 2\lambda^2 c/\beta$ (see (2)). As a result, combining the above and (11) with (10) completes the proof.

References

- [1] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [2] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *IEEE Elec. Lett.*, vol. 27, no. 23, pp. 2145–2146, 1991.
- [3] M. Haenggi, "Diversity polynomials for the analysis of temporal correlations in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5940–5951, 2013.
- [4] R. Tanbourgi, H. Jkel, and F. K. Jondral, "Cooperative relaying in a poisson field of interferers: A diversity order analysis," In *Proc. IEEE ISIT*, Jul. 2013, pp. 3100–3104.
- [5] R. Tanbourgi, H. S. Dhillon, J. G. Andrews, and F. K. Jondral, "Effect of spatial interference correlation on the performance of maximum ratio combining," *IEEE Trans. on Wireless Commun.*, Vol. 13, No. 6, pp. 3307–3316, 2014.
- [6] S. Krishnan and H. S. Dhillon, "Spatio-temporal interference correlation and joint coverage in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5659–5672, 2017.
- [7] T. Kimura and H. Saito, "Temporal correlation of interference under spatially correlated shadowing," In *Proc. WiOpt*, May 2018.
- [8] R. Ganti and M. Haenggi, "Spatial and temporal correlation of the interference in ALOHA ad hoc networks," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 631–633, 2009.
- [9] U. Schilcher, C. Bettstetter, and G. Brandner, "Temporal correlation of interference in wireless networks with Rayleigh block fading," *IEEE Trans. Mobile Comput.*, vol. 11, no. 12, pp. 2109–2120, 2012.
- [10] Z. Gong and M. Haenggi, "Interference and outage in mobile random networks: expectation, distribution, and correlation," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 337–349, 2014.
- [11] J. Wen, M. Sheng, B. Liang, X. Wang, Y. Zhang, and J. Li, "Correlations of interference and link successes in heterogeneous cellular networks," In *Proc. Globecom*, Dec. 2015, pp. 1–6.
- [12] J. Wen, M. Sheng, K. Huang, and J. Li, "Analysis of interference correlation in non-Poisson networks," In *Proc. Globecom*, Dec. 2016, pp. 1–6.
- [13] K. Koufos and C. P. Dettmann, "Temporal correlation of interference in bounded mobile ad hoc networks with blockage," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2494–2497, 2016.
- [14] K. Koufos, C. P. Dettmann, and J. P. Coon, "Correlated interference from uncorrelated users in bounded ad hoc networks with blockage," *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 114–117, 2017.
- [15] K. Koufos, and C. P. Dettmann, "Temporal correlation of interference and outage in mobile networks over one-dimensional finite regions," *IEEE Trans on Mobile Comput.*, vol. 17, no. 2, pp. 475–487, 2018.
- [16] B. Błaszczyszyn and M. K. Karray, "Quality of service in wireless cellular networks subject to log-normal shadowing," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 781–791, 2013.
- [17] B. Błaszczyszyn and M. K. Karray, and H. P. Keeler, "Using Poisson processes to model lattice cellular networks," In *Proc. INFOCOM*, April 2013, pp. 773–781.
- [18] R. W. Heath, Jr., M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using Poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114–4126, 2013.
- [19] M. D. Renzo, A. Guidotti, and G. E. Corazza, "Average rate of downlink heterogeneous cellular networks over generalized fading channels: a stochastic geometry approach," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3050–3071, 2013.
- [20] F. Baccelli and X. Zhang, "A correlated shadowing model for urban wireless networks," In *Proc. INFOCOM*, May 2015, pp. 801–809.
- [21] J. Lee, X. Zhang, and F. Baccelli, "Shadowing and coverage in Poisson buildings," In *Proc. INFOCOM*, April 2016, pp. 1–9.
- [22] J. D. Coninck, F. Dunlop and T. Huillet, "On the correlation structure of some random point processes on the line," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 5, pp. 725–744, 2008.
- [23] M. Haenggi, *Stochastic Geometry for Wireless Networks*, 1st Ed., Cambridge University Press, New York, 2013.
- [24] P. D. Miller, *Applied Asymptotic Analysis*, RI: American Mathematical Society, Providence, 2006.
- [25] T. Kimura and H. Saito, "Spatio-temporal correlation of interference in MANET under spatially correlated shadowing environment," *submitted*, 2018.

SMT ソルバを用いた確率ペトリネットのパーフェクトサンプリングアルゴリズム

岡村 寛之 森原 和也 土肥 正
広島大学大学院工学研究科

概要： 本稿では、確率ペトリネットに対するパーフェクトサンプリングについての考察を行う。パーフェクトサンプリングとは厳密に定常状態における状態をサンプリングする手法である。具体的に本稿では与えられた確率ペトリネットにエンベロープを用いた CFTP (Coupling From The Past) 法を適用するためのアルゴリズムについて提案する。

1. はじめに

システムの信頼性や保全性を定量的に評価するためには確率モデルが必要不可欠である。一般的に、確率モデルを用いたシステムの定量評価ではシステムの振る舞いを状態遷移モデルで表現し、遷移時間がなんらかの確率分布によって推移することを仮定する。よく行われるモデル化にマルコフ連鎖を用いたものがある [1]。マルコフ連鎖とは、ある状態から別の状態へ推移する確率が過去の履歴に関係なく現在の状態のみに依存することを仮定した離散状態の確率過程であり、動作対象により連続時間あるいは離散時間でのシステム動作を表現することができる。

マルコフ連鎖を利用する利点として解析が容易に行えることが挙げられる。システムの定量評価に必要な解析として、無限回の推移後にどの状態を取るのかと言った定常解析とある決められた期間の推移でどの状態を取るのかと言った過渡解析がある。マルコフ連鎖を用いたモデル化では定常解析、過渡解析いずれも解析手法が確立されており、状態数が少ないものであれば比較的簡単にシステムの性能評価尺度を導出することができる。システムの性能評価尺度の導出する手法の一つとしてモンテカルロシミュレーションがある。モンテカルロシミュレーションは疑似乱数を用いてシステムの状態遷移を模倣することで性能評価尺度を見積もる手法である。モンテカルロシミュレーションの利点は多数の状態をもつシステムに対しても適用することができる。しかしながら、高精度の評価を行うためには多数のサンプルが必要になることと、定常解析で得られるような尺度を得るときに「どのくらいの時間推移させたら定常になるのか」が明確ではない。

後者の問題を解決する一つの手法としてパーフェクトサンプリングが提案されている。状態遷移モデルにおけるパーフェクトサンプリングとは確実に定常状態からのサンプルを得るための手法である。具体的に Propp and Wilson [2] は CFTP (Coupling From The Past) 法と呼ばれる手法を提案し、離散時間マルコフ連鎖に対するパーフェクトサンプリングを実現している。特に文献 [2] では単調なマルコフ連鎖に対して単調 CFTP と呼ばれる、より効率的な手法を提案している。一方、対象とするマルコフ連鎖が単調であることは稀であり、一般的なシステムをモデル化した場合、非単調であることが多い。Busic et al. [3] は、非単調なマルコフ連鎖にも適用できるエンベロープ法と呼ばれる CFTP 法の拡張を提案している。

本稿では確率ペトリネットに対するエンベロープ法の適用を考える。実際にエンベロープ法を適用する際には対象となる確率モデルの特徴を加味したアルゴリズムを具体的に構築する必要がある。また、エンベロープ法では順序の定義が任意であるため、定義した順序によっても構築される具体的なアルゴリズムが変化する。これはエンベロープ法を利用するユーザ自身がアルゴリズムをアドホック的に構築する必要性を示唆しており、実際の利用にとって大きな障害となる。そのため、より汎用的な確率モデルに対するエンベロープ法を対象としたアルゴリズムの構築を自動化（ユーザからの隠蔽化）が重要となる。確率ペトリネットは任意の連続時間マルコフ連鎖を表現するコンパクトな表現形式であり、確率ペトリネットに対するエンベロープ法の構築を行うことでアルゴリズム構築の自動化（ユーザからの隠蔽化）を行う。

本稿の構成は以下の通りである。第 2 章では、CFTP 法およびエンベロープ法の基本的な考え方について紹介する。第 3 章では、確率ペトリネットの基本的な性質の紹介と、確率ペトリネット上でエンベロープ法を実現する手続きについて提案する。第 4 章では、確率ペトリネットで記述した待ち

行列ネットワークシステムに対して提案手法を適用し、有効性の検証を行う。第 5 章では、まとめと今後の課題について述べる。

2. パーフェクトサンプリング

2.1. CFTP 法

離散時間の確率過程 $\{X_n, n = 0, 1, 2, \dots\}$ を有限な状態空間 $\mathcal{S}$ 上を推移する既約な離散時間マルコフ連鎖とし、その状態遷移確率行列を $\mathbf{P}$ とする。このとき、定常状態における各状態の確率を表す定常確率（行）ベクトル $\boldsymbol{\pi}$ は以下を満たすベクトルとして得られる。

$$\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}, \quad \boldsymbol{\pi} \mathbf{1} = 1. \quad (1)$$

ここで $\mathbf{1}$ はすべての要素が 1 の列ベクトルである。有限で既約なマルコフ連鎖の場合、上記の定常確率ベクトルはマルコフ連鎖を無限回推移したときの状態に対する分布（極限推移確率）と等しい。

いま離散時間マルコフ連鎖に対するシミュレーションを考えるため更新関数を導入する。 $\mathcal{E}$ をマルコフ連鎖上で発生し得るすべてのイベントの集合とする。このとき、確率過程 X_n の 1 ステップでの推移を表す更新関数 $\Phi(s, e)$, $s \in \mathcal{S}$, $e \in \mathcal{E}$ を定義する。定義から明らかに状態 s_i から状態 s_j 推移するすべてのイベントに対する確率 $\sum_{e \in \mathcal{E}} P(\Phi(s_i, e) = s_j)$ は $\mathbf{P}$ の (i, j) 要素 ($p_{i,j}$ と表記) で表される。

イベント列 $e^{1 \rightarrow n} = \{e^1, e^2, \dots, e^n\}$ に対して次の n ステップ更新関数を定義する。

$$\Phi^n(x, e^{1 \rightarrow n}) = \Phi(\dots \Phi(\Phi(x, e^1), e^2) \dots, e^n). \quad (2)$$

さらに、状態空間の部分集合 $\mathcal{A} \subset \mathcal{S}$ に対する n ステップ更新関数を次のように定義する。

$$\Phi^n(\mathcal{A}, e^{1 \rightarrow n}) = \{\Phi^n(x, e^{1 \rightarrow n}), x \in \mathcal{A}\}. \quad (3)$$

CFTP (Coupling From The Past) 法によるマルコフ連鎖のパーフェクトサンプリングは次の定理をもとにしている [2]。

定理 1 任意の過去 ($t = -n$) から現在 ($t = 0$) までのイベント列 $e^{-n \rightarrow 0}$ に対して、次の式を満たす整数 l が存在する。

$$\lim_{n \rightarrow \infty} |\Phi^n(\mathcal{S}, e^{-n \rightarrow 0})| = l. \text{ (almost surely)} \quad (4)$$

上記の定理は、過去のどのような状態から推移を始めたとしても過去から現在までのイベント列 $e^{-n \rightarrow 0}$ のもとでは現在時刻の（取り得る）状態数が l となることを意味している。つまり、 $l = 1$ となるようなイベント列が生成された場合、過去をどれだけさかのぼったとしても現在の状態は一つに決まるため、それを定常状態（無限回推移した時の状態）とすることができ。つまり、CFTP 法ではすべての状態から始まる状態推移が一つの状態になる（収斂 (coalesce) と呼ぶ) まで過去にさかのぼってイベントを生成し、現在時点で収斂しているマルコフ連鎖の状態を定常分布に対するサンプルとして採用する。

しかしながら、実際のシミュレーションでは例えば状態空間が有限であってもマルコフ連鎖が取り得るすべての状態を予め求めておくことは難しい。また、仮にすべての状態がわかったとしても、すべての状態から始まるマルコフ連鎖のシミュレーションをすることは非常に計算コストがかかる。そこで、CFTP 法に対していくつかの工夫が必要となる。文献 [2] では、マルコフ連鎖の単調性を利用した単調 CFTP 法を提案している。マルコフ連鎖の単調性とは、マルコフ連鎖の状態空間 $\mathcal{S}$ がある半順序 $\preceq$ で順序付けられているものとする、すべてのイベント $e \in \mathcal{E}$ に対して

$$\forall (x, y) \in \mathcal{S}, x \preceq y \Rightarrow \Phi(x, e) \preceq \Phi(y, e). \quad (5)$$

が成り立つことを表している。これは、同じイベントならば、推移を行っても推移後の順序が変化しないことを意味している。更新関数の意味での単調性は確率順序に基づくマルコフ連鎖の単調性と同値であることが示されている。単調 CFTP 法では、状態空間の最大元と最小元に対応する状態からのマルコフ連鎖シミュレーションのみを考える。上記の単調性の性質から、すべての状態は最大元と最小元から始めたマルコフ連鎖の間に存在することが保証されるため、最大元、最小元からのマルコフ連鎖が現在時刻で同じ値を取ったならば、現在時刻のマルコフ連鎖の状態は収斂しており、それを定常分布に対するサンプルとすることができる。

2.2. エンベロープ法

文献 [3] では単調でないマルコフ連鎖に対する CFTP 法としてエンベロープ (Envelope) を用いた手法を提案している。単調 CFTP 法と同様に、マルコフ連鎖の状態空間 $\mathcal{S}$ に対して半順序 $\preceq$ を定義する。さらに、状態空間 $\mathcal{S}$ を含む完備束 (complete lattice) $\mathcal{S}^*$ を考える。

次に、拡張した状態空間の直積 $\mathcal{S}^* \times \mathcal{S}^*$ とイベント集合 $\mathcal{E}$ に対するエンベロープ更新関数 Γ を考える。

$$(m', M') = \Gamma(m, M, e), \quad m, M, m', M' \in \mathcal{S}^*, e \in \mathcal{E}. \quad (6)$$

ここで $\Gamma(m, M, e)$ はもとの更新関数を用いて次のように定義される。

$$\Gamma(m, M, e) = (\bar{\Phi}(m, M, e), \Phi(m, M, e)), \quad (7)$$

$$\Phi(m, M, e) = \inf_{m \preceq x \preceq M, x \in \mathcal{S}} \Phi(x, e), \quad (8)$$

$$\bar{\Phi}(m, M, e) = \sup_{m \preceq x \preceq M, x \in \mathcal{S}} \Phi(x, e). \quad (9)$$

ここで、 $\Phi(m, M, e)$ と $\bar{\Phi}(m, M, e)$ は状態 m, M の間に挟まれたマルコフ連鎖が実際に取り得る状態すべてをイベント e によって更新した新しい状態に対する下限と上限である。いま、 $\mathcal{S}^*$ は完備束と仮定しているためそのような下限・上限は唯一に決定することに注意する。さらに、イベント列 $e^{1 \rightarrow n}$ に対する n ステップエンベロープ更新関数を

$$\Gamma^n(m, M, e^{1 \rightarrow n}) = \Gamma(\cdots \Gamma(\Gamma(m, M, e^1), e^2) \cdots, e^n) \quad (10)$$

と定義する。このとき、上記の n ステップエンベロープ更新関数でシミュレートされる $(m, M), \Gamma(m, M, e^1), \dots$ の列は状態空間 $\mathcal{S}^* \times \mathcal{S}^*$ 上のマルコフ連鎖となる。エンベロープ法は次の定理にもとづいている [3]。

定理 2 任意の過去 ($t = -n$) から現在 ($t = 0$) までのイベント列 $e^{-n \rightarrow 0}$ に対して、次の式を満たす τ が有限に存在するものと仮定する。

$$\tau = \min \{n : \Gamma^n(B, T, e^{-n \rightarrow 0}) \in \mathcal{D}\}. \quad (11)$$

ここで $\mathcal{D} = \{(m, M); m = M, m, M \in \mathcal{S}^*\}$ 。また、 $B = \inf \mathcal{S}^*$ および $T = \sup \mathcal{S}^*$ である。このとき、 τ は $\mathcal{S}$ 上のもとのマルコフ連鎖 X_n の後ろ向きカップリング時間 (収斂時間) となる。

上記の定理は、エンベロープ更新関数 Γ によって推移する状態が現在時刻において $m = M$ となっている時、その状態がもとのマルコフ連鎖の状態空間 $\mathcal{S}$ の要素であり、それをもとのマルコフ連鎖の定常状態に対するサンプルとして採択して良いことを示している。直感的な説明としては、もとのマルコフ連鎖の全状態を挟む上下限を (m, M) で表現し、上下限だけに着目したマルコフ連鎖を考えることで単調でないマルコフ連鎖に対する CFTP を実現している。しかしながら、エンベロープ更新関数の作成方法は任意であり、必ずしも $m = M$ とならないことに注意する。例えば、すべての (m, M) に対して

$$\Phi(m, M, e) = B, \quad \bar{\Phi}(m, M, e) = T \quad (12)$$

と定義されるエンベロープ更新関数は要件を満たしているが明らかに $m = M$ とはなり得ない。従って、より狭い上下限を与えるエンベロープ更新関数の作成方法が重要になる。文献 [4] では、より狭い上下限を与える手法として、エンベロープ法では順序の定義は任意であることを利用した異なる順序空間に基づいた多点のエンベロープ更新関数を議論している。

3. 確率ペトリネットに対するエンベロープ法

3.1. 確率ペトリネット

ペトリネット (Petri Net; PN) はプレース (Place)、トランジション (Transition)、トークン (Token) で構成され、プレースとトランジションが接続する二部有向グラフで表現される。あるトランジションに注目したとき、そのトランジションに向かう矢印に接続しているプレースを入力プレース、トラ

ランジションから出る矢印に接続しているプレースを出力プレースと呼ぶ。トークンはプレースに配置される。すべての入力プレースにトークンが存在する時、そのランジションは発火可能 (Enable) といい、すべての発火可能なランジションのうち、いずれか一つが発火 (Firing) する。ランジションの発火によってプレースに配置されたトークン数が変化する。あるランジションが発火した場合、そのランジションの入力プレースに配置されたトークンが一つずつ取り除かれ、出力プレースにトークンが一つずつ配置される。

PN では、全プレースに対するトークンの配置状態 (マーキング) によってシステムの状態を表現する。ランジションの発火によるマーキングの推移をマーキングプロセスと呼び、これが状態遷移モデルに対応する。またモデルの表現力を向上させるため、ランジションの発火に条件を課して、その条件が満たされない場合は発火を抑止することが行われる。この条件をガード条件 (Guard Condition) と呼び、ガード条件を許す PN の表現力はチューリング機械と同等であることが知られている [5]。

ランジションが発火可能となってから発火するまでの具体的な時間 (発火遅延) を考慮したペトリネットを時間ペトリネット (Timed PN) と呼び、その発火遅延が指数分布に従うランジション (指数ランジション) で構成されるものを確率ペトリネット (Stochastic PN; SPN) と呼ぶ。これは状態遷移時間がすべて指数分布で記述されるため、SPN のマーキングプロセスは連続時間マルコフ連鎖に帰着される。

3.2. 状態方程式とインバリアント

PN のプレース数を N_P , ランジション数を N_T とする。このとき、 $N_T \times N_P$ の接続行列 $\mathbf{A} = [a_{i,j}]$ を定義する。ここで $a_{i,j}$ はランジション j が発火した時、プレース i で起きるトークン数の変化量を表している。例えば、 $a_{1,5} = -1$ であれば、ランジション 5 が発火した時、プレース 1 におけるトークン数が一つ減ることを意味している。PN において現在のマーキングを表す列ベクトル (各プレースに配置されたトークン数の列ベクトル) を $\mathbf{x}$ とし、次の発火を表す発火ベクトルを $\mathbf{r}$ とする。発火ベクトルは発火するランジションを 1, それ以外のランジションを 0 とする列ベクトルである。このとき、次の発火後のマーキングベクトル $\mathbf{x}'$ は接続行列を用いて

$$\mathbf{x}' = \mathbf{x} + \mathbf{A}\mathbf{r} \quad (13)$$

として表現される。上記を PN の状態方程式と呼ぶ。

いま、 $\mathbf{A}^T \mathbf{J} = \mathbf{0}$ を満たす非負整数解 $\mathbf{J}$ が存在するものとする。ここで T は行列の転置を表す。このとき、 $\mathbf{J}$ は P -インバリアントと呼ばれ、状態方程式から直ちに

$$\mathbf{J}^T \mathbf{x} = \mathbf{J}^T \mathbf{x}' \quad (14)$$

となることがわかる。一般に、マーキング $\mathbf{x}$ が初期マーキング $\mathbf{x}_0$ から推移したマーキングであるための必要条件は P -インバリアント $\mathbf{J}$ を用いて

$$\mathbf{J}^T \mathbf{x}_0 = \mathbf{J}^T \mathbf{x} \quad (15)$$

となる。

3.3. イベント列の定義

SPN のマーキングが構成する連続時間マルコフ連鎖にパーフェクトサンプリングを適用するために離散時間マルコフ連鎖を構成する必要がある。いま、SPN のマーキングが構成する連続時間マルコフ連鎖の無限小生成行列を $\mathbf{Q}$ とすると、無限小生成行列の対角要素 $q_{i,i}$ の絶対値よりも大きい $q > \max |q_{i,i}|$ を考え、次の変換を行うことを一様化と言う。

$$\mathbf{P} = \mathbf{I} + \mathbf{Q}/q. \quad (16)$$

ここで、 $\mathbf{I}$ は単位行列である。このようにして得られた $\mathbf{P}$ は離散時間マルコフ連鎖の推移確率行列となり、 $\mathbf{P}$ の定常分布はもとの連続時間マルコフ連鎖の定常分布と一致することが知られている。

本稿では、SPN における指数ランジションの発火レートはマーキングに依存しないものとし、 q を全ての指数ランジションの発火レートの和で与える。このとき、明らかに $q > \max |q_{i,i}|$ を満た

す。また、変換された離散時間マルコフ連鎖では、イベント列として「どのトランジションが発火するか」というイベントを選ぶことができる。具体的に「 j 番目のトランジションが発火する」というイベント e_j は次の確率によって選ばれる。

$$P(e_j) = \frac{\lambda_j}{\sum_{i=1}^{N_T} \lambda_i}. \quad (17)$$

ここで λ_j は j 番目の指数トランジションに対する発火レートである。

3.4. 上下限値の導出

先に示したようにエンベロープ法では更新関数として、与えられた上下限値 (m, M) であるイベント e が発生した後の上下限値 (m', M') を導出する必要がある。いま、SPN のマーキングベクトル $\mathbf{x}$ に対して次のような半順序を考える。

$$\begin{aligned} \mathbf{x} = (x_1, x_2, \dots, x_{N_P}) \preceq \mathbf{y} = (y_1, y_2, \dots, y_{N_P}) \\ \text{if } x_i \leq y_i \text{ } i = 1, 2, \dots, N_P. \end{aligned} \quad (18)$$

上記の半順序は本質的ではなく、後の説明を簡単にするために設定している。順序の関係は 3.6 で拡張される。

いま、イベントとして e_j (トランジション j が発火する) が選択されたものとする。このとき、イベント発生後の上下限値を求める問題は、以下の最大/最小化問題を解くことに帰着される。

Solve

$$m'_i = \min x'_i \quad \text{for } i = 1, \dots, N_P, \quad (19)$$

$$M'_i = \max x'_i \quad \text{for } i = 1, \dots, N_P, \quad (20)$$

Subject to

$$\text{enable}_j(\mathbf{x}) \rightarrow \mathbf{x}' = \mathbf{x} + [\mathbf{A}]_j, \quad (21)$$

$$\overline{\text{enable}_j(\mathbf{x})} \rightarrow \mathbf{x}' = \mathbf{x}, \quad (22)$$

$$m_i \leq x_i \leq M_i \quad \text{for } i = 1, \dots, N_P, \quad (23)$$

$$\mathbf{J}^T \mathbf{x}_0 = \mathbf{J}^T \mathbf{x}. \quad (24)$$

ここで、 m_i および M_i は上下限値 (m, M) における各要素を表す。また、 $\text{enable}_j(x_1, \dots, x_n)$ は状態ベクトルが $\mathbf{x} = (x_1, \dots, x_n)$ の時にトランジション j が発火可能かどうかを表す論理値、 $\rightarrow$ は論理包含、 $\bar{a}$ は論理値 a の否定を表す。さらに $[\mathbf{A}]_j$ は接続行列の j 列目のベクトルを表す。エンベロープ法では、SPN の状態ベクトルが実際に取り得るものだけを考える必要があるため、3.2 で示した P -インバリアントによる必要条件を加えている。

3.5. SMT ソルバの利用

前述の上下限値の導出はトランジションがガード条件を含まない場合や、 P -インバリアントが単純なものについては具体的に解くことができる。しかしながら、トランジションがガード条件を含むような状況においては容易に解くことができない。そこで、本稿では SMT (Satisfiability Modulo Theories) ソルバを用いたアルゴリズムを提案する。具体的に状態ベクトルの上下限が (m, M) でイベント e_j (トランジション j が発火) した場合に、新たな上下限 (m', M') を求める手続きを Algorithm 1 と Algorithm 2 に示す。例えば Algorithm 1 では m' の各要素について可能性のある範囲で小さな値から順番に条件を満たすかどうか判定し、最初に条件を満たした数値で m' の要素を与えている。

3.6. 状態および順序の拡張

ここまでの議論では、SPN の状態をマーキングベクトルで与え、各プレースのトークン数による順序を用いた。ここでは、これらの拡張について述べる。

Algorithm 1 A function to obtain the lower bound vector for the event e_j ; $\Phi(m, M, e_j)$.

```

1: for  $i = 1$  to  $N_P$  do
2:   for  $m'_i = \min(m_i, m_i + a_{i,j})$  to  $\max(M_i, M_i + a_{i,j})$  do
3:      $s \leftarrow$  solve the satisfiability problem with SMT solver:

```

$$\begin{aligned}
& m'_i = x'_i, \\
& \text{enable}_j(\mathbf{x}) \rightarrow \mathbf{x}' = \mathbf{x} + [\mathbf{A}]_j, \\
& \overline{\text{enable}_j(\mathbf{x})} \rightarrow \mathbf{x}' = \mathbf{x}, \\
& m_i \leq x_i \leq M_i \quad \text{for } i = 1, \dots, N_P, \\
& \mathbf{J}^T \mathbf{x}_0 = \mathbf{J}^T \mathbf{x}.
\end{aligned}$$

```

4:   if  $s = \text{true}$  then {The problem is satisfied.}
5:     break
6:   end if
7: end for
8: end for
9: return  $m' = (m'_1, \dots, m'_{N_P})$ 

```

SPN のマーキング集合 $\mathcal{M}$ が有限であると仮定する. さらに, 整数値の集合 $\mathcal{Z}$ を用いて単射 $f: \mathcal{M} \rightarrow \mathcal{Z}^n$ を定義する. このとき, PN のマーキング $\mathbf{x}$ を f を用いて整数値のベクトルに変換して表現することができる.

$$\mathbf{z} = (z_1, z_2, \dots, z_n) = f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{M}. \quad (25)$$

さらに, 上記の整数値のベクトルに対して次のような半順序を与える.

$$\begin{aligned}
\mathbf{z} = (z_1, z_2, \dots, z_n) \preceq \mathbf{z}' = (z'_1, z'_2, \dots, z'_n) \\
\text{if } z_i \leq z'_i \text{ } i = 1, 2, \dots, n.
\end{aligned} \quad (26)$$

例えば, 上記の拡張において z_i をプレース i に配置されているトークン数に -1 を乗じたものと定義することもでき, この場合は全てのプレースでトークン数が多い方が小さい順序 (逆順) となる.

また, すべての要素が整数からなる $n \times N_P$ の行列 $\mathbf{B}$ で次のような線形写像を考えることもできる.

$$\mathbf{z} = \mathbf{B}\mathbf{x}. \quad (27)$$

この場合, 上下限値を求めるための問題は以下のようなになる.

Solve

$$m'_i = \min z'_i \quad \text{for } i = 1, \dots, n, \quad (28)$$

$$M'_i = \max z'_i \quad \text{for } i = 1, \dots, n, \quad (29)$$

Subject to

$$\text{enable}_j(\mathbf{x}) \rightarrow \mathbf{z}' = \mathbf{z} + [\mathbf{B}\mathbf{A}]_j, \quad (30)$$

$$\overline{\text{enable}_j(\mathbf{x})} \rightarrow \mathbf{z}' = \mathbf{z}, \quad (31)$$

$$m_i \leq z_i \leq M_i \quad \text{for } i = 1, \dots, n, \quad (32)$$

$$\mathbf{J}^T \mathbf{x}_0 = \mathbf{J}^T \mathbf{x}, \quad (33)$$

$$\mathbf{z} = \mathbf{B}\mathbf{x}. \quad (34)$$

Algorithm 2 A function to obtain the upper bound vector for the event e_j ; $\overline{\Phi}(m, M, e_j)$.

```

1: for  $i = 1$  to  $N_P$  do
2:   for  $M'_i = \max(M_i, M_i + a_{i,j})$  to  $\min(m_i, m_i + a_{i,j})$  do
3:      $s \leftarrow$  solve the satisfiability problem with SMT solver:

```

$$\begin{aligned}
M'_i &= x'_i, \\
\text{enable}_j(\mathbf{x}) \rightarrow \mathbf{x}' &= \mathbf{x} + [\mathbf{A}]_j, \\
\overline{\text{enable}_j(\mathbf{x})} \rightarrow \mathbf{x}' &= \mathbf{x}, \\
m_i \leq x_i \leq M_i &\quad \text{for } i = 1, \dots, N_P, \\
\mathbf{J}^T \mathbf{x}_0 &= \mathbf{J}^T \mathbf{x}.
\end{aligned}$$

```

4:   if  $s = \text{true}$  then {The problem is satisfied.}
5:     break
6:   end if
7: end for
8: end for
9: return  $M' = (M'_1, \dots, M'_{N_P})$ 

```

4. 適用例

ここでは、6 ステーションからなる開放型の待ち行列ネットワークで SPN に対するエンベロープ法の適用例を示す。図 1 は待ち行列ネットワークを表す SPN であり、各プレイス $p1$ から $p6$ に配置されるトークン数が各ステーションの客数に対応する。ここでは各ステーションに収容できる客数が有限であり、最大収容客数を K で表している。そのため、トランジション $t1$ から $t5$ については最大収容数を超えないようにガード条件が付与されている。また、厳密な意味での P -インバリアントは存在しないが、トークンの配置は以下の条件を満たす必要がある。

$$p_2 + p_4 = p_3 + p_5.$$

ここで p_i はプレイス i に配置されているトークン数を表す。

いま、状態ベクトルを $\mathbf{x} = (p_1, p_2, p_3, p_4, p_5, p_6)$ と定義する。SPN を一様化して得られるマルコフ連鎖は単調ではないためエンベロープ法の適用を行う必要がある。一方、すべてのプレイスでトークン数が 0 となる状態が最小元、すべてのプレイスでトークン数が K となる場合が最大元となる。つまり、 $m = (0, 0, 0, 0, 0, 0)$ と $M = (K, K, K, K, K, K)$ からあるイベント列に対してエンベロープ法を適用することとなる。しかしながら、上下限値が $m = (*, 0, 0, 0, 0, *)$ と $M = (*, K, K, K, K, *)$ となっている場合（プレイス $p1$ と $p6$ の上下限値は任意）、以下の考察からトランジション $t1$ から $t6$ のいずれが発火するイベントが選ばれても、プレイス $p2$ から $p5$ までの上下限値に変化はないことがわかる。

- トランジション $t1$ と $t6$ については、発火したとしてもプレイス $p2$ から $p5$ のトークン数に影響を与えないため、 $p2$ から $p5$ の上限値は変わらない。
- トランジション $t3$ と $t4$ について $\mathbf{x} = (*, K, K, K, K, *)$ および $\mathbf{x} = (*, 0, 0, 0, 0, *)$ とすると発火できない（ガード条件を満たさないあるいは入力プレイスのトークン数が 0）ため、 $p2$ から $p5$ の上下限値は変わらない。
- トランジション $t2$ と $t5$ について $\mathbf{x} = (*, 0, K, K, 0, *)$ および $\mathbf{x} = (*, K, 0, 0, K, *)$ とすると発火できない（ガード条件を満たさないあるいは入力プレイスの一方のトークン数が 0）ため、 $p2$ から $p5$ までの上下限値は変わらない。

つまり、この状態ベクトルではエンベロープ法を適用したとしても収斂しない。

上述の問題があるため本稿では状態ベクトルを $\mathbf{x} = (p_1, p_2, p_3, p_4, p_5, p_6, \sum_{i=1}^6 p_i)$ と定義し、最小元 $m = (0, 0, 0, 0, 0, 0, 0)$ 、最大元 $M = (K, K, K, K, K, K, 6K)$ としてエンベロープ法を実施する。

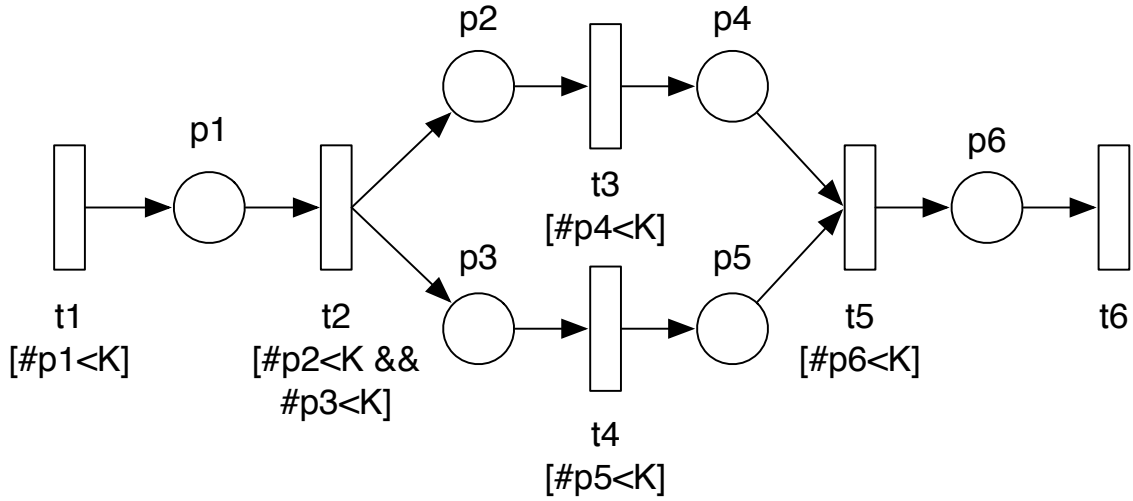


図 1: SPN.

これは 3.6 において

$$B = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (35)$$

とすることに対応する．SPN 内のトークン数の合計を入れた拡張された状態空間を使うことで前述したような状況がなくなりエンベロップ法が収斂する．

図 2 は $K = 4$ とし，各トランジションの発火率を 1 に設定したとき，定常状態でプレース p2 と p6 が $p2=4$, $p6=4$ となる確率をエンベロップ法により推定した結果を示している．ここでは，エンベロップ法から得た 1000 サンプルから定常状態確率を推定することとし，それを 10 回繰り返したときのばらつきを箱ひげ図で示している．また，図上の横線は解析的に得られた定常状態確率を表している．これらの図から，1000 サンプルからの推定はある程度ばらつくものの，10 回の平均値は実際の解析的に得られた定常状態確率に非常に近くなっており，提案した SPN に対するエンベロップ法により定常状態からのサンプルが得られていることが確認できる．

5. まとめと今後の課題

本稿では SPN に対するパーフェクトサンプリングについて考察した．ここでは，SMT ソルバを使うことで，どのような SPN に対しても汎用的に適用できるエンベロップ法アルゴリズムを開発した．また，具体的な SPN に適用し定常分布に従うサンプルが取得できることを確認した．

ここで提案したアルゴリズムは汎用的であるものの，現時点では適用できる SPN のクラスについての限定が行えていない．つまり， P -インバリアントによる必要条件で得られるマーキング集合と実際に SPN が取り得るマーキング集合が同じもので検証しているが，これが異なる場合にも提案したエンベロップ法が機能するかどうかについての考察が行えていない．また，適応例ではすべてのプレースに配置されたトークン数を補助的に状態ベクトルに組み入れることでエンベロップ法が機能することがわかったが，どのような状態を補助的に入れる必要があるかについての考察も十分に行えていない．今後は数理的な観点からそれらの条件を明確に示す必要がある．

一方，アルゴリズムでは SMT ソルバを利用しているため，サンプルを生成するまでの計算時間が多くかかる．具体的に適応例であげた SPN では 1 つのサンプルを得るのに平均で 57.72 秒の時間がかかっている．このため，この部分の高速化が課題となっている．さらに，より適用範囲を広げるた

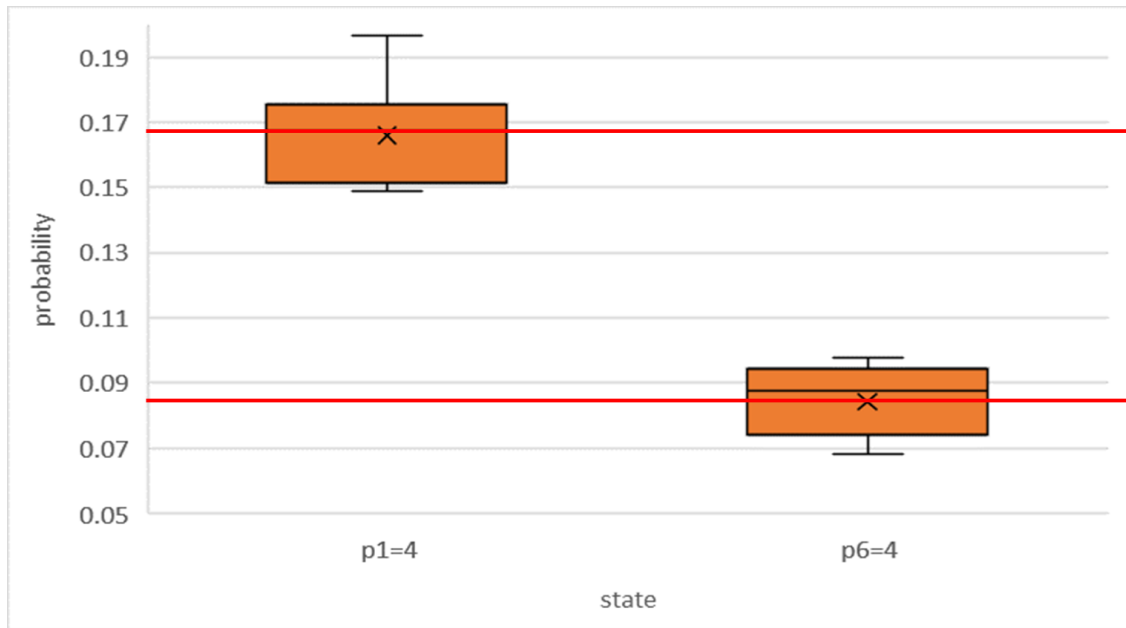


図 2: 定常状態における確率の推定.

めには瞬時トランジションを含む一般化確率ペトリネット (Generalized SPN; GSPN) について適用できるように拡張する必要がある. また, もとの SPN が吸収状態を持つ場合にはエンベロープ法そのものが機能しなくなるため, その点についての改善も必要となる.

参考文献

- [1] K. S. Trivedi. *Probability and Statistics with Reliability, Queueing, and Computer Sciences Applications*. John Wiley & Sons, New York, 2nd edition, 2001.
- [2] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9(1-2):223–252, 1996.
- [3] A. Busic, B. Gaujal, and J.-M. Vincent. Perfect simulation and non-monotone Markovian systems. In *Proceedings of 3rd International Conference on Performance Evaluation Methodologies and Tools (Valuetools 2008)*, pages 1–10, 2008.
- [4] 岡村 寛之, 森原 和也, 土肥 正. 位相型近似による非マルコフモデルのパーフェクトサンプリングに関する一考察. 信学技報, R2018-9, pages 13–19, 2018.
- [5] 村田 忠夫. ペトリネットの解析と応用. 近代科学社, 1992.

ハウスドルフ一次モーメント型トラフィックモデルの性能評価に対する不完全性

高田 寛之, 長崎大学

概要： 本稿では極端に条件が少ないトラフィックモデルを用いて待ち行列モデルの性能評価を行う。このトラフィックモデルは確率論のモーメント問題と呼ばれる分類名に従って著者が勝手に作った名称であるが、その実体はサポートが有界でその境界及び期待値の上界が既知な確率変数列である。ここでは、平均待ち行列長の期待値の上界評価を数値的に行い、単位時間を小さくすると、その上界はほとんど最悪値と同じ値になってしまう現象について述べる。タイトルの不完全性が指し示すのはこの現象のことである。また、なぜそのような集中現象が起きるのかを確率のサポートの状況を見ることで説明する。

1. はじめに

1.1. 何を調べるか

本稿では、連続時間待ち行列モデルの平均待ち行列長の上界評価を行う。具体的に $\sup_{A \in \mathcal{T}_1(r,b)} E[Q(t)]$ の値に興味がある。ここでそれぞれの記号は以下のように定義される。 $0 \leq u \leq t$ となる u, t に対して、 $A(u, t)$ は区間 $(u, t]$ におけるシステムへの累積入力量を表す。システムは入力関数 A とサービス率 $c > 0$ の仕事保存サーバによって動作が決まる。待ち行列長は $Q(t) = \sup_{0 \leq u \leq t} A(u, t) - c * (t - u)$ で定義される。 $r > 0, b > 0$ はそれぞれトラフィックの平均流量の上界と最大バースト値を表す。安定条件 $r < c$ を仮定する。考慮するトラフィックモデルは次のようなクラスである。

$$\mathcal{A}_1(r, b) = \{A \mid \Pr[0 \leq A(u, t) \leq r(t - u) + b] = 1, \quad E[A(u, t)] \leq r(t - u), \quad (0 \leq u \leq t)\}.$$

1.2. 何故調べるか

ここでは、なぜこんなモデルを調べるのかの理由を述べる。トラフィック特徴を単純化することで計測オーバーヘッドを削減することができるかもしれない。また仮定成立を保証するためのトラフィック制御装置の開発が容易になるかもしれない。

単純化することの影響は上界の緩さとして反映される。どの程度悪いのか。かなり悪いならなぜ悪くなるのかその仕組みを知って、非自明な（タイトな）上界評価に必要な最低限のモデルは何かをつきとめたい。そのために最も単純だと思われる期待値モデルを考えている。

条件を追加していくことで必要最小限の仮定を構築できるものと信じている。

1.3. キーアイデア

サポートが有界でその境界と期待値の上界が与えられた条件の下で、最大順序統計量の期待値を目的関数として最大化問題を解いた結果 [1] を適用することで、このモデルの上界評価の限界を明らかにする。ただし、問題は連続時間モデルであり、この結果は離散時間モデルであるので、離散時間近似してから適用する。離散時間近似に伴う極限操作は、有界収束定理により期待値の外に出せる。

高田 [1] の結果を簡単に述べる。ベクトル間の不等式は項別不等式を表し、確率ベクトルの期待値は項別期待値のベクトルを表す。 $\xi_0 \leq \xi_a \leq \xi_1$ はそれぞれ n 次元実ベクトルとする。 $\xi_{i,j}$ はベクトル ξ_i の第 j 要素を表す。 $\xi_{1,0} = \max_{i=1}^n \xi_{0,i}$ とおく。 $\xi_{1,0} > \xi_{1,i}$ なる項 i がある場合は、予めその項を除去しておく。更に $\xi_{1,1} \leq \xi_{1,2} \leq \dots \leq \xi_{1,n}$ であるとする。このようにして $\xi_{1,0} \leq \xi_{1,1} \leq \dots \leq \xi_{1,n}$ であるとする。 $x^+ = \max(0, x)$ とする。

補題 1 (高田 [1]).

$$\sup \left\{ E \left[\max_{i=1}^n X_i \right] \mid \Pr(\xi_0 \leq \mathbf{X} \leq \xi_1) = 1, \quad E[\mathbf{X}] \leq \xi_a \right\} = E[Y]$$

が成り立つ。ここで、 Y は次のように定義される離散型確率変数である。

$$\Pr[Y = \xi_{1,j}] = q_{j+1}^+ - q_j^+ \quad (j = 0, \dots, n),$$
$$q_j = \begin{cases} 0 & j = 0 \\ 1 - \sum_{k=j}^n p_k & j = 1, \dots, n \end{cases}, \quad p_k = \begin{cases} \frac{\xi_{a,k} - \xi_{0,k}}{\xi_{1,k} - \xi_{0,k}} & \xi_{1,k} > \xi_{0,k} \\ 1 & \xi_{1,k} = \xi_{0,k} \end{cases} \quad (k = 1, \dots, n).$$

離散近似の式を以下に述べる。

補題 2. $f(x)$ が右連続ならば、

$$\sup_{0 \leq x \leq x_0} f(x) = \lim_{n \rightarrow \infty} \max_{k=0}^n f\left(\frac{x_0 * k}{n}\right).$$

2. 結果

補題 1 と補題 2 を用いて、 $\sup_{A \in \mathcal{A}_1(r,b)} E[Q(t)]$ を計算すると、

$$\sup_{A \in \mathcal{A}_1(r,b)} E[Q(t)] = \lim_{n \rightarrow \infty} \sum_{j=0}^{\tau(an)} b \left(1 - \frac{j}{n}\right) \frac{j}{j+an},$$

$$a = \frac{c-r}{r}, \quad \tau(x) = \max \left\{ n \in \mathbb{N} \left| \sum_{k=1}^n \frac{k}{k+x} < 1 \right. \right\}.$$

ここで、 $\tau(x)$ は大体 $\sqrt{2x}$ で近似できる。

定理 1. 任意の正の実数 x に対して $\tau(x)$ は $\lfloor \sqrt{2x} \rfloor$ か $\lceil \sqrt{2x} \rceil$ のどちらかに等しい。

従って、平均待ち行列長の上界は、 $\lim_{n \rightarrow \infty} \sum_{j=0}^{\text{int}(\sqrt{2na})} b \left(1 - \frac{j}{n}\right) \frac{j}{j+an} = b$ と計算される。つまり、確率変数 Y は n が十分大きいとき、そのサポートが $(b(1 - \sqrt{2a/n}), b)$ に集中する。従ってほぼ最悪値評価の結果である $Q(t) \leq b$ と同じ評価しか得られず、統計的多重効果をとることができない。

一方、 n が十分小さいときは、この上界は b より小さい値をとる。これは、一見不思議な結果であるが、よく考えれば、1 分単位で道路が渋滞していても 1 日単位では道路は渋滞していない（1 日単位で観測して道路が渋滞するということは相当酷い渋滞がずっと続いている）と考えれば、自然な結果である。実際の通信ではシステムの駆動単位時間はかなり小さいので n が十分大きい値のときの評価が適切である。

3. まとめと相談ごと

本報告では、連続時間待ち行列モデルにおいて、ハウスドルフ一次モーメントトラフィックモデルを仮定して平均待ち行列長の上界評価を行った。結果として、最悪値評価と変わらず、ハウスドルフ一次モーメントトラフィックモデルでは条件が不十分であることがわかった。従って、一次モーメント以外の確率的条件を追加する必要がある。

みなさんに相談したいのは、次に調査すべきトラフィックモデルの条件を決めるために次のような問題ではどのように合格者数を決めるかをお伺いしたい。定員が 100 名の私立大学入試を行う。ここで合格は一回でまとめて出すこととする。定員の 80% 以上を充足する必要があるが、進学者数は 120 名を超えてはいけない。過去の辞退者数の履歴データを見ることはできる。（もしこのような処理をしたことがある人がいれば、どういう理屈で合格者数を決定するのかを知りたい）もしくはこれに類似する問題の解決例があればお伺いしたい。なお、多変量モーメント問題はモーメントの情報が 2 次以上のケースは比較的論文が多い。

参考文献

- [1] H. Takada, “Cross moments approaches for mean-bounded finite random classes,” in *the Queue Symposium 2017*, jan 2018.

Queues driven by level dependent Lévy processes

Irmina Czarna José-Luis Pérez Tomasz Rolski Kazutoshi Yamazaki
 Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology
 Department of Probability and Statistics, Centro de Investigación en Matemáticas
 Mathematical Institute, University of Wrocław
 Department of Mathematics, Kansai University

Abstract Motivated by the M/G/1 queues with abandonment, we study a generalization of the Lévy queue whose drift term is decreasing in the queue level. This is modeled by the stochastic differential equation (SDE) of the form

$$dU(t) = dX(t) - \phi(U(t)) dt, \quad t \geq 0,$$

for a given spectrally negative Lévy process X and an increasing function ϕ . We analyze the solutions to the above SDE and investigate the so-called *scale functions*, which are counterparts of the scale functions from the theory of Lévy processes. We show how fluctuation identities for U can be expressed via these scale functions.

1. Introduction

We consider continuous-time queueing models driven by Lévy processes with negative jumps (spectrally negative Lévy processes). Our motivation is to model queues with abandonment where a customer is not patient enough to line up if the queue is long, and consequently the rate of increments of a queue decreases when the queue length is high.

We start with an ordinary spectrally negative Lévy process $X = \{X(t), t \geq 0\}$ with its Laplace exponent:

$$\psi(\lambda) := \log \left[e^{\lambda X(t)} \right] = \gamma\lambda + \frac{1}{2}\sigma^2\lambda^2 + \int_0^\infty \left(e^{-\lambda z} - 1 + \lambda z \mathbf{1}_{(0,1]}(z) \right) \Pi(dz), \quad \lambda \geq 0,$$

for $\gamma \in \mathbb{R}$, $\sigma \geq 0$, and a σ -finite measure Π on $(0, \infty)$ such that $\int_0^\infty (1 \wedge z^2) \Pi(dz) < \infty$. Our objective is to study the evolution of a queue U as a solution to the SDE

$$dU(t) = dX(t) - \phi(U(t)) dt, \quad t \geq 0. \tag{1}$$

Here, the degree of abandonment is modeled by the measurable function ϕ , which is assumed to be non-decreasing.

As a special case of the SDE (1), Kyprianou and Loeffen [1] studied the *refracted Lévy process*, which is the unique strong solution to the SDE (1) for the case $\phi(x) = \delta \mathbf{1}_{\{x > b\}}$ with $\delta > 0$. Using the scale functions $W^{(q)}$ and $Z^{(q)}$ of the process X (see [1]) and those for the drift-changed process $\{X(t) - \delta t, t \geq 0\}$, they computed various expectations of interest.

2. Multi-refracted Lévy processes

We first extend the refracted Lévy process [1] to a multi-refracted case with

$$\phi(x) = \phi_k(x) = \sum_{i=1}^k \delta_i \mathbf{1}_{\{x > b_i\}}, \tag{2}$$

for $\delta_1, \dots, \delta_k > 0$ and refraction levels $b_1 < \dots < b_k$. For the case X is of bounded variation, we assume additionally that

$$\phi(x) < \gamma + \int_0^1 z \Pi(dz), \quad x \in \mathbb{R}, \tag{3}$$

so that the process does not have monotone paths on any interval.

Under these conditions, there exists a unique strong solution U_k to the SDE (1), and we call it the *multi-refracted Lévy process*. For $a \in \mathbb{R}$, let

$$\kappa_k^{a,-} := \inf\{t > 0: U_k(t) < a\} \quad \text{and} \quad \kappa_k^{a,+} := \inf\{t > 0: U_k(t) \geq a\},$$

with the convention that $\inf \emptyset = \infty$. Our first main results are given as follows.

Theorem 1. *Using the scale functions $w_k^{(q)}$ and $z_k^{(q)}$ as defined in Section 2 of [2], we have the following expressions:*

1. *Two-sided exit identities:*

$$\begin{aligned} \mathbb{E}_x \left[e^{-q\kappa_k^{a,+}} \mathbf{1}_{\{\kappa_k^{a,+} < \kappa_k^{d,-}\}} \right] &= \frac{w_k^{(q)}(x; d)}{w_k^{(q)}(a; d)}, \\ \mathbb{E}_x \left[e^{-q\kappa_k^{0,-}} \mathbf{1}_{\{\kappa_k^{0,-} < \kappa_k^{a,+}\}} \right] &= z_k^{(q)}(x) - \frac{z_k^{(q)}(a)}{w_k^{(q)}(a; 0)} w_k^{(q)}(x; 0). \end{aligned}$$

2. *Resolvents:*

$$\mathbb{E}_x \left[\int_0^{\kappa_k^{a,+} \wedge \kappa_k^{d,-}} e^{-qt} \mathbf{1}_{\{U_k(t) \in \mathcal{B}\}} dt \right] = \int_{\mathcal{B} \cap (d, a)} \frac{\frac{w_k^{(q)}(x; d)}{w_k^{(q)}(a; d)} w_k^{(q)}(a; y) - w_k^{(q)}(x; y)}{1 - W^{(q)}(0) \phi_k(y)} dy.$$

3. Level dependent Lévy processes

We now consider the SDE with ϕ such that it is non-decreasing and continuously differentiable and $\phi(x) = 0$ for $x \leq 0$. We additionally impose (3) for the case X is of bounded variation. By taking a sequence of $(\phi_n)_{n \geq 0}$ of the form (2) converging to ϕ , we analyze the solution to (1) and obtain its fluctuation identities.

Under these conditions, there indeed exists a unique strong solution U . By taking limits on the results in Theorem 1, we have our second main results.

Theorem 2. *For $a \in \mathbb{R}$, let $\kappa^{a,-} := \inf\{t > 0: U(t) < a\}$ and $\kappa^{a,+} := \inf\{t > 0: U(t) \geq a\}$. Using the scale function $w^{(q)}$ and $z^{(q)}$ as defined in Section 3 of [2], we have the following:*

1. *Two-sided exit identities:*

$$\begin{aligned} \mathbb{E}_x \left[e^{-q\kappa^{a,+}} \mathbf{1}_{\{\kappa^{a,+} < \kappa^{d,-}\}} \right] &= \frac{w^{(q)}(x; d)}{w^{(q)}(a; d)}, \\ \mathbb{E}_x \left[e^{-q\kappa^{0,-}} \mathbf{1}_{\{\kappa^{0,-} < \kappa^{a,+}\}} \right] &= z^{(q)}(x) - \frac{z^{(q)}(a)}{w^{(q)}(a; 0)} w^{(q)}(x; 0). \end{aligned}$$

2. *Resolvents:*

$$\mathbb{E}_x \left[\int_0^{\kappa^{a,+} \wedge \kappa^{d,-}} e^{-qt} \mathbf{1}_{\{U(t) \in \mathcal{B}\}} dt \right] = \int_{\mathcal{B} \cap (d, a)} \frac{\frac{w^{(q)}(x; d)}{w^{(q)}(a; d)} w^{(q)}(a; y) - w^{(q)}(x; y)}{1 - W^{(q)}(0) \phi(y)} dy.$$

References

- [1] KYPRIANOU, A.E., LOEFFEN, R. Refracted Lévy processes. *Ann. Inst. H. Poincaré*, 46 (1), 24–44, (2010).
- [2] CZARNA, I., PÉREZ, J-L, ROLSKI, T, YAMAZAKI, K. Fluctuation theory for Level-dependent Lévy risk processes. *arXiv:1712.00050*, (2017).

真社会性昆虫の基本再生産数

石井 諒太郎 豊泉 洋
早稲田大学 基幹理工学部 応用数理学科

1. はじめに

真社会性とは次の 3 つの条件：(1) 子の保護における協同, (2) 繁殖に関する分業, すなわち多少とも不妊の個体が繁殖個体を助けて働くこと, (3) コロニーの労働が可能な発育段階の少なくとも 2 世代にわたる共存, をすべて満たす集団であり, その集団の中で他の個体の利益のために行動する個体が存在する [1]. 真社会性昆虫の繁殖戦略はマルコフ連鎖や分枝過程を用いて研究されている [2],[3]. 真社会性昆虫の繁殖戦略 ([2] は巣の相続がない場合,[3] は巣の相続がある場合) を比較する評価基準として基本再生産数が導出されている. この場合における基本再生産数とは, ある 1 つの巣について発生してから死滅するまでに新しく発生させる巣の数の期待値である.

今回は [2],[3] で導出された基本再生産数を相型分布と基本的な線形代数の行列演算を用いて導出できることを示す.

2. 巣の中の個体数のマルコフ連鎖と相型分布表現

巣の中の個体数を巣の大きさとし, その最大を M とする. 巣の大きさ i における昆虫の出生レート, 死亡レートをそれぞれ b_i, d_i ($i = 1, \dots, M$), 巣に残る確率を q とする. また, 大きさ i のときの巣の大きさの増加レートと巣の発生レートを

$$\mu_i = \begin{cases} qb_i & \text{if } 1 \leq i \leq M-1 \\ 0 & \text{if } i = M \end{cases}, \quad \nu_i = \begin{cases} (1-q)b_i & \text{if } 1 \leq i \leq M-1 \\ b_M & \text{if } i = M \end{cases}, \quad \nu = (\nu_1, \dots, \nu_M)^T$$

とおく. 時刻 t で巣の大きさが i である確率を $\pi_i(t)$ とし, $\pi(t) = (\pi_1(t), \pi_2(t), \dots, \pi_M(t))$, $\pi(0) = (1, 0, \dots, 0)$ とする. $\{\pi(t); t \geq 0\}$ の時間変化は連続時間マルコフ連鎖に従うとする. [2],[3] の各場合における推移率行列 $Q = (Q_{ij})_{M \times M}$, $Q' = (Q'_{ij})_{M \times M}$ はそれぞれ

$$Q_{ij} = \begin{cases} -(\mu_i + d_i) & \text{if } i = j \\ \mu_i & \text{if } j = i + 1 \\ 0 & \text{otherwise} \end{cases}, \quad Q'_{ij} = \begin{cases} -(\mu_i + d_i) & \text{if } i = j \\ \mu_i & \text{if } j = i + 1 \\ d_i & \text{if } i = j + 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

であり, いずれも 0 は吸収状態となる. また [2],[3] の各場合において $(0, \pi(0))$ から始めて巣の大きさが 0 になるまでの時間 T, T' はそれぞれ表現 $(\pi(0), Q)$, 表現 $(\pi(0), Q')$ の相型分布に従う.

$\{\pi(t); t \geq 0\}$ は過渡状態の連続時間マルコフ連鎖に従い, 巣の相続がない場合, $\frac{d\pi}{dt}(t) = \pi(t)Q$ が成立し, この微分方程式の解は $\pi(t) = \pi(0) \exp(Qt)$ となる. T は表現 $(\pi(0), Q)$ の相型分布に従い, $\det Q = (-1)^M \prod_{i=1}^M (\mu_i + d_i) \neq 0$ で Q は正則であるからそれぞれの巣の大きさから確率 1 で吸収状態である巣の大きさ 0 になり, $\int_0^\infty \exp(Qt) dt = -Q^{-1}$ となる ([4] Theorem 2.4.3). したがって基本再生産数 $E[R]$ は,

$$E[R] = \int_0^\infty \pi(t) \nu dt = \int_0^\infty \pi(0) \exp(Qt) \nu dt = \pi(0)(-Q^{-1}) \nu \quad (2)$$

となる. 巣の相続がある場合も同様に $E[R] = \pi(0)(-Q'^{-1}) \nu$ で基本再生産数を求められる.

2.1. 巣の相続がない場合の基本再生産数

[2] では真社会性昆虫 ($M \geq 1$) の基本再生産数 $E[R]$ を行列の第 1 固有値, 行列のスペクトル半径を用いて,

$$E[R] = \sum_{i=1}^M \frac{\nu_i}{\mu_i + d_i} \prod_{j=1}^{i-1} \frac{\mu_j}{\mu_j + d_j} \quad (3)$$

と導出している.

Q の i 行 j 列を除いた小行列式 $\tilde{q}_{ij}$ は小行列を区分けすることにより,

$$\tilde{q}_{ij} = \begin{cases} (-1)^{M-1} \left(\prod_{k=1}^{j-1} (\mu_k + d_k) \right) \left(\prod_{k=j}^{i-1} \mu_k \right) \left(\prod_{k=i+1}^M (\mu_k + d_k) \right) & \text{if } i \geq j \\ 0 & \text{otherwise} \end{cases}$$

となり, Q の余因子行列 $\tilde{Q} = (\tilde{q}_{ji})_{M \times M}$ を求められる. ゆえに逆行列 $Q^{-1} = (Q_{ij}^{-1})_{M \times M}$ は,

$$Q_{ij}^{-1} = \begin{cases} -\prod_{k=i}^{j-1} \mu_k / \prod_{k=i}^j (\mu_k + d_k) & \text{if } i \leq j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

と求められる. (2) 式, (4) 式より (3) 式が成立することがわかる.

2.2. 巣の相続がある場合の基本再生産数

[3] では真社会性昆虫の基本再生産数 $E[R]$ を指数分布の無記憶性を用いた確率的な議論により,

$$E[R] = \sum_{i=1}^M \frac{\nu_i}{d_i} \prod_{j=1}^{i-1} \frac{\mu_j}{d_j} \quad (5)$$

と導出している.

Q' の余因子行列の各成分を求めるうえで, 任意の自然数 $n \leq M$ に対して $n \times n$ 行列 $T^{(n)}$

$$T_{ij}^{(n)} = \begin{cases} -(\mu_i + d_i) & \text{if } i = j \\ \mu_i & \text{if } j = i + 1 \\ d_i & \text{if } i = j + 1 \\ 0 & \text{otherwise} \end{cases}$$

を考え, その行列式 $\det T^{(n)}$ を考える. 漸化式 $\det T^{(n+1)} = -\mu_n \det T^{(n)} + (-1)^{n+1} \prod_{l=1}^{n+1} d_l$ ($n \geq 1$) が成立し,

$$\det T^{(n)} = (-1)^n \sum_{k=0}^n \left(\prod_{l=1}^{n-k} d_l \right) \left(\prod_{l=n-k+1}^n \mu_l \right) \quad (n \geq 1) \quad (6)$$

と解くことができる. Q' の i 行 j 列を除いた小行列式 $\tilde{q}'_{ij}$ は小行列を区分けし, (6) 式を用いると,

$$\tilde{q}'_{ij} = \begin{cases} (-1)^{M-i} \left(\prod_{l=i+1}^M d_l \right) \det T^{(i-1)} & \text{if } i > j \\ (-1)^{M-j} \left(\prod_{l=i+1}^M d_l \right) \left(\prod_{l=j}^{i-1} \mu_l \right) \det T^{(j-1)} & \text{otherwise} \end{cases}$$

となり, Q' の余因子行列 $\tilde{Q}' = (\tilde{q}'_{ji})_{M \times M}$ を求められる. ゆえに逆行列 $Q'^{-1} = (Q'_{ij})^{-1}_{M \times M}$ は

$$Q'_{ij}^{-1} = \begin{cases} -\left\{ \sum_{k=0}^{i-1} \left(\prod_{l=1}^{i-k-1} d_l \right) \left(\prod_{l=i-k}^{j-1} \mu_l \right) \right\} / \left(\prod_{l=1}^j d_l \right) & \text{if } i < j \\ -\left\{ \sum_{k=0}^{j-1} \left(\prod_{l=1}^{j-k-1} d_l \right) \left(\prod_{l=j-k}^{i-1} \mu_l \right) \right\} / \left(\prod_{l=1}^j d_l \right) & \text{otherwise} \end{cases} \quad (7)$$

と求められる. (2) 式, (7) 式より (5) 式が成立することがわかる.

参考文献

- [1] 伊藤 嘉昭. 動物の社会 社会生物学・行動生態学入門. 東海大学出版会. (2006)
- [2] Feng Fu, Sarah D. Kocher and Martin A. Nowak. The risk-return trade-off between solitary and eusocial reproduction. *Ecology Letters* (2015) 18:74-84
- [3] Hiroshi Toyoizumi and Jeremy Field. A Markov Model of Eusociality at its Origin. 2018 年度待ち行列シンポジウム. (2019)
- [4] G. Latouche, V. Ramaswami. Introduction to Matrix Analytic Methods in Stochastic Modeling. SIAM, Philadelphia. (1999)

乗車時間を考慮したタクシー待ち行列モデルの解析

巻野侑大

Phung-Duc Tuan

筑波大学理工学群社会工学類

筑波大学システム情報系

1. はじめに

中長距離を移動する様々な交通手段が存在する現代において、タクシーによる移動は、目的地や出発地点を選ぶことができる便利な移動手段の1つである。タクシーサービスが展開されているタクシースタンドでは、乗客とタクシーの双方の待ち行列が構成される。本研究では、駅や空港などのターミナル施設にあるタクシースポットで展開されるタクシーサービスの過程を、待ち行列モデルとしてモデル化し、その解析を行う。

本研究で想定しているような、顧客とサービスを提供する媒体の2つの待ち行列が存在する過程をモデル化し解析している研究には、サービス時間を考慮していないモデルと、考慮したモデルの2種類がある。前者の研究[1]では、2重待ち行列モデルとして過程をモデル化し、解析している。また、後者の研究[2]では、マッチングを行う窓口を1つとしてモデル化し、その解析を行っている。実際のタクシーサービスの現場では、乗車しスタンドから出発するまでに時間がかかることが考えられ、また駅や空港などにあるターミナル施設では、スタンドにある複数の乗降口からサービスが展開されている。そのため、サービス時間を考慮し、窓口（乗降口）数を複数個としたモデル化を行うことにより、従来のサービス時間を考慮していないモデルや、窓口が1つのモデルと比べ、より現実のタクシースポットの状況に近いモデルの解析を行うことができる。

本研究では、サービス時間を考慮し、かつ複数の窓口を用いたタクシーサービスの過程をモデル化し、その性能指標を定義・導出する。最後に数値実験を行い、結果について考察を行う。

2. モデル化と解析

本研究では、サービス時間を考慮し、また複数個の窓口があるタクシースタンドにてサービスが行われることを考える。乗客とタクシーは、それぞれポアソン過程に従うパラメータ λ_p, λ_T でスタンドに到着する。到着した乗客は、乗車中でないタクシーが系内におり、かつ乗車中のタクシーで窓口がふさがっていない場合、タクシーとマッチングする。スタンドに到着したタクシーについても、乗車していない乗客が系内におり、かつ窓口がふさがっていない場合、乗客とマッチングする。マッチングした乗客とタクシーにかかるサービス（乗車）時間は、パラメータ μ の指数分布に従い、乗車完了後系内を離れる。系内にマッチングできる乗客またはタクシーがない、またはサービス中ですべての窓口がふさがっている場合、乗客またはタクシーは系内にて待つ。また、系内に駐車できるタクシー数を K とし、系内の窓口の数を c とする。以上より、今回想定するタクシーサービスモデルは図1のような待ち行列モデルになる。

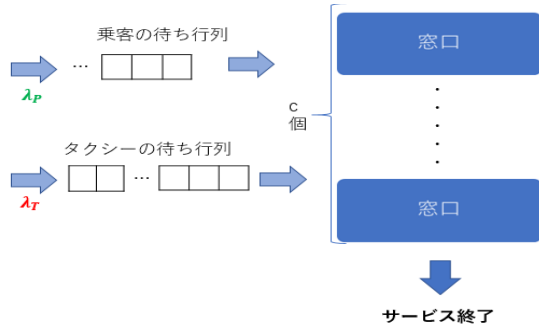


図 1：タクシースタンドの待ち行列.

ここで、集合 I, J をそれぞれ

$$I = \{0, 1, 2, \dots\}, \quad J = \{0, 1, 2, \dots, K-1, K\},$$

で定義し、時刻 t における系内の乗客の数を $i(t)$ 、タクシーの数を $j(t)$ とすると、

$$(i(t), j(t)) \in S = \{(i, j) \mid i \in I, j \in J\}.$$

となり、 $\{(i(t), j(t)); t \geq 0\}$ は S 上のマルコフ連鎖になる.

また、サービス時間は窓口の状況によって異なるため、状態 (i, j) におけるサービス率を以下のように定義する.

$$\mu_{i,j} = \min(i\mu, j\mu, c\mu).$$

次に平衡方程式を立て、定常状態確率の計算を行う. $\{(i(t), j(t)); t \geq 0\}$ は準出生死滅過程であるため、定常分布の計算は[3]で示された方法を用いた.

また、 $e = (1, 1, \dots, 1)'$ 、状態 (i, j) における定常状態確率を $\pi_{i,j}$ 、 $\pi_i = (\pi_{i,0}, \pi_{i,1}, \dots, \pi_{i,K})$ とし、平均系内乗客数を $E[L_P]$ 、平均系内タクシー数を $E[L_T]$ 、乗客の平均滞在時間を $E[W_P]$ 、タクシーの平均滞在時間を $E[W_T]$ とすると、本研究で用いる各性能指標は以下のように表すことができる.

- ・ 平均系内乗客数 $E[L_P]$: $\sum_{i=0}^{\infty} i \cdot \pi_i \cdot e$.
- ・ 平均系内タクシー数 $E[L_T]$: $\sum_{i=0}^{\infty} \sum_{j=0}^K j \cdot \pi_{i,j}$.
- ・ 乗客が待つ確率 : $1 - \sum_{i=0}^c \sum_{j=i}^K \pi_{i,j}$.
- ・ タクシーが待つ確率 : $\sum_{i=0}^c \sum_{j=i+1}^K \pi_{i,j} + \sum_{i=c+1}^{\infty} \sum_{j=c+1}^K \pi_{i,j}$.
- ・ 乗客の平均滞在時間 $E[W_P]$: $E[L_P]/\lambda_P$.
- ・ タクシーの平均滞在時間 $E[W_T]$: $E[L_T]/\lambda_T$.

参考文献

- [1] 松島格也・小林潔司・坂口潤一 (2003) タクシースポット市場の差別化と社会的厚生, 土木学会論文集, 2003(723), pp. 41-53.
- [2] Jascha Prosiel (2012) An Assembly Queue with Infinite and Finite Waiting Spaces, 東京工業大学・交換留学プログラム最終レポート, pp. 2-16.
- [3] 滝根哲哉 (2014) M/M/1 を越えて—準出生死滅過程への招待—, オペレーションズリサーチ, 59(4), pp. 3-8.

閉園時刻直前のアトラクションの選び方

鶴田 裕太郎 佐藤 真史 豊泉 洋
早稲田大学大学院 基幹理工学研究科

概要：閉園時刻直前で、その時点での待ち人数と各アトラクションへの到着レートが既知の場合に、閉園前に乗るべきアトラクションの判定法について述べる。さらに、この判定法では、自分より後に到着した客が理性的に動くと、乗れるはずのアトラクションに乗れない悲劇が起こる事例を紹介する。

1. はじめに

テーマパークにはアトラクションが多数存在する。待ち時間が長く混雑しているテーマパークでは、客は一つでも多くのアトラクションに乗る事を望んでいる。テーマパークの問題解決について過去に研究がなされている。パーク側の観点から、優先パスを用いることで平均待ち時間を減らし混雑度を緩和するという研究が行われている [1]。客の動きを考えたゲーム理論的待ち行列は [2] で詳しく考察されているが、今回は、一定間隔で客が到着する場合に、多くアトラクションに乗る最適な回り方についてモデルする。例として2つのアトラクション「トイストーリーマニア」と「タワーオブテラー」について閉園時刻直前で両方乗ることができるか、それとも片方のみ乗ることができるかを判定する方法を示し、さらに、このモデル化で自分より後に到着した客の自由な動きを考慮した場合、悲劇が起こってしまう事例を紹介する。

2. アトラクションモデル

人気度の違うアトラクション1,2を考える。アトラクション i の待機人数を l_i 、客の予想到着レートを λ_i 、閉園までの残り時間を T とした時に到着した客の戦略(アトラクションに乗る順番)を考える。客は、アトラクションの列に並んでいる人数 (l_1, l_2) を観測可能で、到着レート (λ_1, λ_2) を予想可能とする。アトラクション i のサービス時間(アトラクションに乗っている時間)は s_i であり、 x_i 人を間隔 i_i ごとに乗せるものとする。列は単位時間あたり $c_i = x_i/i_i$ 人進む。

3. (λ_1, λ_2) が固定されている場合の戦略

客が等間隔($1/\lambda_i = \text{const}$)で来るとする。閉園時刻の T 分前に来た客(ターゲット)を考える。客の持つ情報を $(l_1, l_2, T, \lambda_1, \lambda_2)$ とする。この時、アトラクション1に乗れる条件は式(1)である。

$$l_1 \leq c_1 T \quad (1)$$

また、アトラクション2に乗った上に、アトラクション1に乗れる条件はアトラクション2の滞在時間 $\frac{l_2}{c_2} + s_2$ と、アトラクション1の待ち列の人数 l_1 より式(2)である。

$$l_1 + \left(\frac{l_2}{c_2} + s_2 \right) \lambda_1 \leq c_1 T \quad (2)$$

今回、閉園1時間前($T=60$)の時($\lambda_1 = 9, \lambda_2 = 4.5$)のトイストーリーマニアとタワーオブテラーで検証を行った。待ち人数の観測結果によるトイストーリーマニアの前にタワーオブテラーへの乗車可能性の判定結果を図1に示す。例えば、 $(l_1, l_2) = (350, 50)$ の時、どちらから先に乗っても両方乗ることができる。また、 $(l_1, l_2) = (440, 5)$ の時は、先にタワーオブテラーの乗ると、両方乗ることができる。

4. 他者の動きを反映させた場合の悲劇

ターゲット以外の客も図1に従って動くものとした場合悲劇が起こる。図1の $(l_1, l_2) = (462, 0)$ から始まった場合の悲劇が起こる状況について図2に示す。ターゲットの客Aは、 $(l_1, l_2, T, \lambda_1, \lambda_2) = (462, 0, 60, 9, 4.5)$ なので、図1の判定により、2から先に乗る。客Aが2に並んだことにより、 l_2 が1人増える。次の客Bは $(l_1, l_2, T, \lambda_1, \lambda_2) = (462, 1, 60 - \frac{1}{13.5}, 9, 4.5)$ なので、式(2)より片方にしか乗れない。

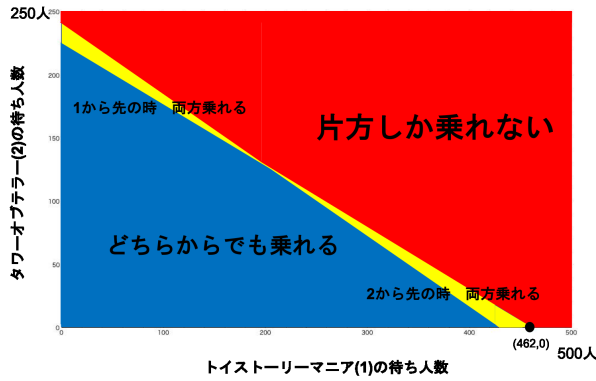


図 1: 予想到着レート $(\lambda_1, \lambda_2) = (9, 4.5)$,
 $(s_1 = 5, i_1 = 1, x_1 = 8, c_1 = 8,$
 $s_2 = 2, i_2 = 5, x_2 = 22, c_2 = 4.4)$ の場合

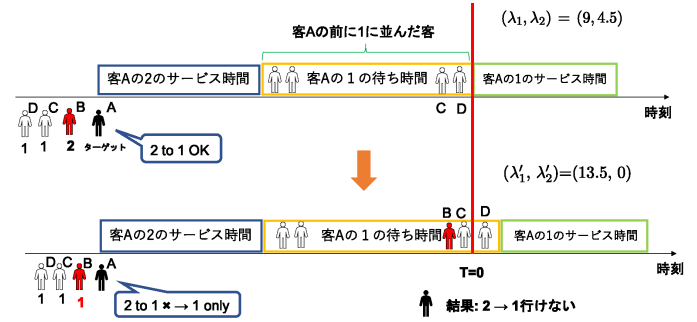


図 2: 悲劇が起こる状況
 (客 B が片方しか乗れず, 魅力度により 1 に乗
 るため客 A が 1 に乗れなくなった)

結果, 客 A は 2 to 1 の順でアトラクションに乗れ, 客 B は片方しか乗れない. 客 B は, より魅力のあるアトラクションを選択し 1 に行く. 結果, 客 A がアトラクション 2 から退出する時刻では, 予想より 1 の列が長くなりターゲットはアトラクション 1 に乗れないという悲劇が起こる. これは, ターゲットの客がアトラクション 2 から退出する時刻 $60 - s_2$ において, $(l_1, l_2, T, \lambda'_1, \lambda'_2) = (462 + 5.5s_2, -4.4s_2, 60 - s_2, 13.5, 0)$ となり, 式 (2) を満たさないからである.

5. 今後の展望

本研究を基礎とし, 確率的動的計画法の最短経路問題 [4] を応用し, アトラクション数を実際のテーマパークと同数にした場合のモデルの構築および解析を行う. 悲劇を防ぐ対策として, 優先パスの有効性を評価をする.

参考文献

- [1] Yasushi Masuda, Akira Tsuji, Congestion Control for a System with Parallel Stations and Homogeneous Customers using Priority Passes, Networks and Spatial Economics, (2018).
- [2] Refael Hassin, Moshe Haviv, Queue or not to queue: Equilibrium behavior in queueing systems, Kluwer Academic Publishers, (2003).
- [3] 東京ディズニーリゾートアトラクション情報 <https://www.tokyodisneyresort.jp/tds/attraction.html> (最終閲覧日 2018 年 12 月 11 日).
- [4] Dimitri P. Bertsekas and John N. Tsitsiklis, Neuro-Dynamic Programming, Athena Scientific, Belmont, Massachusetts, (2016).

木構造を持つ一製品多工程拡張かんばんシステムの解析

島先 孝太郎

東京工業大学 情報理工学院

概要：プル型生産システムは需要の発生によって後工程が前工程から部品を引き取って生産を行う生産システムである。前工程は後工程に引き取られて初めて生産を行うため、需要の変動に対応しやすく仕掛品の削減が行いやすい。本研究ではプル型生産システムの一つである拡張かんばんシステムを扱う。離散時間待ち行列システムを用いて木構造を持つ単一品種多工程拡張かんばんシステムのモデル化を行い、そのモデルが M/G/1 型離散時間マルコフ連鎖に帰着できることを示す。

1. はじめに

サプライチェーンとは、一つの企業の生産システムだけでなく、複数の企業間の連携による原料調達から生産管理、物流、販売、配送までの総合的な生産・物流システムのことを指す。サプライチェーンの運用でコストの削減を考える際に、全体の仕掛品（未完成の製品）や在庫の削減と販売機会の増大の両立が求められる。しかし販売機会の増大のためには在庫量を増加する必要がある。このトレードオフに関して、在庫数を最適化するために考えられたのがプル型生産システム（プル生産方式、引っ張り生産方式）[1] である。

需要の予測に基づいて需要の発生の前に生産が開始される従来のプッシュ型生産システムとは異なり、プル型生産システムは実際の需要に応じて需要の発生した後に生産が開始される。現在までに考案されたプル型生産システムとして、かんばん方式 [1]、基点在庫方式 [2]、CONWIP[3]、ハイブリッド方式 [4]、拡張かんばんシステム [5] などがある。かんばん方式は、「かんばん」と呼ばれる情報伝達ツールをもちいて各工程の部品・製品の在庫量を制御し、システム全体の仕掛品を抑制する。基点在庫方式はエシェロン在庫というその工程を含めた下流全ての製品在庫量をもちいて、各工程にそれぞれ基準量（基点在庫量）をあらかじめ定めておいて、エシェロン在庫量が基点在庫量を下回ったら基点在庫量まで生産するという規則によってシステム全体の製品在庫量を抑制する。CONWIP（CONstant Work In Process）は生産システム全体の全仕掛品（Work In Process）を常に一定量に定めるシステムである。ハイブリッド方式、拡張かんばんシステムはそれぞれかんばん方式と CONWIP、かんばん方式と基点在庫方式の混成方式である。これら 5 つの方式で比較するとシミュレーションを元にした性能比較にて在庫の適切な調整という観点で拡張かんばんシステムが最も優秀なシステムであることが示されている [6]。

拡張かんばんシステムがかんばん方式と基点在庫方式の混成方式であることはすでに述べたが、混成によって両方式の長所を引き継ぎ欠点を補い合っている。かんばん方式は、部品・製品の在庫の制御に強い効果を発揮するが、顧客からの需要情報が全工程に伝わるまでに時間がかかる欠点があった。しかし拡張かんばんシステムでは、基点在庫方式の特徴である需要情報が発生と同時に全工程に伝わる仕組みを持っているのでかんばん方式の欠点を補っている。また、基点在庫方式では個々の工程の製品在庫を抑制することができたが、全体での在庫の抑制は苦手としていた。これに対して拡張かんばんシステムでは、かんばん方式の特徴である引き取りかんばんの導入によって、全体での部品・製品の両在庫量を抑制することができる。

本研究では多工程の拡張かんばんシステムの解析を目標としている。かんばん方式について Kimura and Terada[7] によって定式化されて以降、数学モデルをもちいた研究がされている。離散時間マルコフ連鎖をもちいた研究として、Kirkavak and Diner[8] が単一品種、単一かんばん、多工程のかんばん方式に対して、スループット、平均在庫量、設備利用率などを解析している。また、Iwase and Ohno[9] は、単一品種、2 かんばん、多工程のかんばん方式を離散時間マルコフ連鎖として定式化し、M/G/1 型マルコフ連鎖として表現できることを示した。そして、拡張かんばんシステムについては、影島、井家、佐久間 [10] が単一品種、単一工程拡張かんばんシステムについて、M/G/1 型マルコフ連鎖として表現できることを示した。

本研究では木構造を持つ単一品種多工程拡張かんばんシステムについて、M/G/1 型マルコフ連鎖

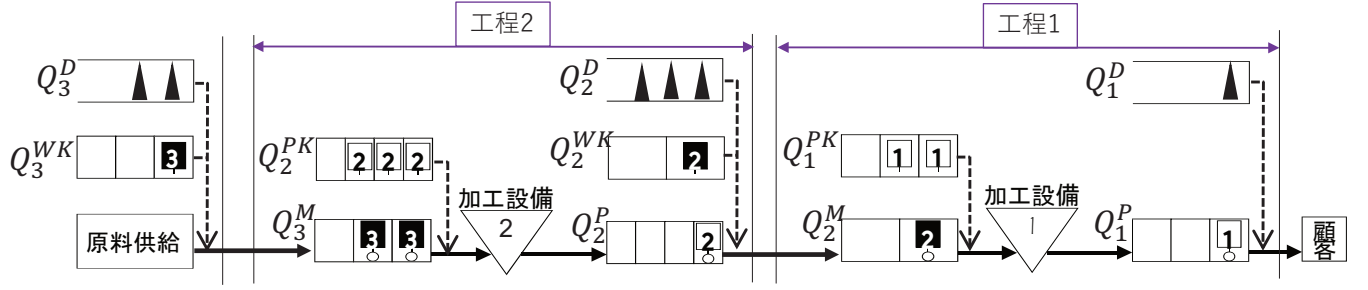


図 1: 2 工程拡張かんばんシステムモデル.

に帰着できることを示す.

2. システムモデルの説明

表 1: 本研究で用いる記号の定義

記号	意味
$\mathbb{N}$	$\{1, 2, \dots\}$
$\mathbb{N}_0$	$\{0, 1, \dots\}$
N	自社工程の数
M	サプライヤーの数
$\mathcal{F}$	$\{1, 2, \dots, N\}$, 自社工程のインデックス全ての集合
$\mathcal{S}$	$\{N+1, N+2, \dots, N+M\}$, サプライヤーのインデックス全ての集合
$\mathcal{P}$	$\mathcal{F} \cup \mathcal{S}$
$\mathcal{P}_{-1}$	$\mathcal{P} \setminus \{1\}$
$\text{pre}(i)$	工程 i の前工程のインデックス全ての集合, $i \in \mathcal{P}$
$\text{pre}(i, j)$	工程 i の前工程の中で j 番目の工程のインデックス, $i \in \mathcal{P}$
$\text{post}(i)$	工程 i の後工程のインデックス, $i \in \mathcal{P}$
i'	$\text{post}(i)$ の略記, $i \in \mathcal{P}$
$1'$	0
$\text{postN}(i)$	工程 i の後工程内でのインデックス, $i \in \mathcal{P}$
$\text{postAll}(i)$	工程 i から後にある工程のインデックス全ての集合, $i \in \mathcal{P}$
n_i	工程 i の前工程の総数, $i \in \mathcal{F}$
Q_i^M	工程 i の部品置き場, $i \in \mathcal{P}_{-1}$
Q_j^P	工程 j の製品置き場, $j \in \mathcal{F}$
Q_i^{WK}	工程 i の引き取りかんばんポスト, $i \in \mathcal{P}_{-1}$
Q_j^{PK}	工程 j の生産指示かんばんポスト, $j \in \mathcal{F}$
Q_k^D	工程 k の受注残, $k \in \mathcal{P}$
S_i^M	工程 i の部品の初期在庫量 (定数), $i \in \mathcal{P}_{-1}$
S_j^P	工程 j の製品の初期在庫量 (定数), $j \in \mathcal{F}$
K_i^W	工程 i の引き取りかんばんの総数 (定数), $i \in \mathcal{P}_{-1}$
K_j^P	工程 j の生産指示かんばんの総数 (定数), $j \in \mathcal{F}$

図 1 のような, M 個のサプライヤーから受け取った部品を N 個の生産工程で加工した後に顧客に 1 種類の製品を納入する, 木構造を持つ単一製品 N 工程拡張かんばんシステムを考える. 簡単のために, 顧客を工程 0, 自社工程を工程 $i \in \mathcal{F} := \{1, 2, \dots, N\}$, (特に最終工程を工程 1), サプライヤーを工程 $j \in \mathcal{S} := \{N+1, N+2, \dots, N+M\}$ とする. サプライヤーは十分な供給能力を持ち, 納入に必要な条件を満たせば次期の期首に直ちに納入する. 生産工程では, 生産に必要な条件が満たされれば直ちに生産し, 部品の引き取り (前工程から後工程へ部品が納入されること) も条件を満たせば次期首に直に行われる. 簡単のために生産工程では前工程から供給された部品全てを 1 単位用いて製品

1 単位を生産すると仮定する。

図1のように、顧客を除いた各工程の構成要素は部品 (Materials) 置き場 Q_i^M , $i \in \mathcal{P}_{-1} := \{2, 3, \dots, N+M\}$, 製品 (Products) 置き場 Q_j^P , $j \in \mathcal{F}$, 引き取り (Withdraw) かんばんポスト Q_i^{WK} , $i \in \mathcal{P}_{-1}$, 生産指示 (Product Instruction) かんばんポスト Q_j^{PK} , $j \in \mathcal{F}$, 需要が到着する受注残 Q_k^D , $k \in \mathcal{P}$, 加工設備である。各工程が持つ在庫について、加工がされる前の在庫を部品、加工がされた後の在庫を製品と呼ぶことにする。

各工程がサプライヤーから部品を受け取る際にサプライヤーは引き取りかんばんポストから引き取りかんばんを取り出し、部品に引き取りかんばんを取り付けて後工程に納入する。後工程は引き取りかんばんを付けたまま部品置き場に部品を置いておき、生産するタイミングで引き取りかんばんを生産指示かんばんに付け替える。外した引き取りかんばんは引き取りかんばんポストに戻す。加工が終わった製品は製品置き場に置かれる。この製品を次の工程に納入するタイミングで生産指示かんばんを引き取りかんばんに付け替える。外した生産指示かんばんは生産指示かんばんポストに戻す。

工程 $i \in \mathcal{P}_{-1}$ の部品の初期在庫量を定数 S_i^M , 工程 $i \in \mathcal{F}$ の製品の初期在庫量を定数 S_i^P , 工程 $i \in \mathcal{P}_{-1}$ の引き取りかんばんの総数を定数 K_i^W , 工程 $i \in \mathcal{F}$ の生産指示かんばんの総数を定数 K_i^P とする。空き引き取りかんばんと需要によって引き取り量が制限され、空き生産指示かんばんによって生産量が制限される。このとき、設定した引き取り指示かんばん枚数が最大部品在庫量であり、設定した生産指示かんばん枚数が最大製品在庫量である。つまり、初期在庫量 $\{S_i^M, S_j^P; i \in \mathcal{P}_{-1}, j \in \mathcal{F}\}$ とかんばん枚数 $\{K_i^M, K_j^P; i \in \mathcal{P}_{-1}, j \in \mathcal{F}\}$ によってシステムの動きを制御できる。ここで $\mathbf{X} = \{S_i^M, S_j^P, K_i^M, K_j^P\} (i \in \mathcal{P}_{-1}, j \in \mathcal{F})$ として、この $\mathbf{X}$ を制御パラメータもしくは初期値と呼ぶことにする。

各期間において、次のような順序でイベントが発生する。

1. サプライヤー、各工程が次工程または顧客に製品を発送する。
2. 各工程、顧客へ製品が到着する。
3. 各工程が製品を生産する。
4. 顧客から各工程へ需要が到着する。

本研究ではシステムの状態観測点をイベント2の終了直後とする。イベント1では製品の発送が行われる。期間 $n \in \mathbb{N}$ での工程 i が発送する製品数を $O_i^S(n)$ とする。イベント2では部品の受取が行われる。期間 n で工程 i が引き取る製品量を $O_i^A(n)$ とする。期間 n の状態観測点における工程 i の部品在庫量を $Q_i^M(n)$, 工程 i の製品の在庫量を $Q_i^P(n)$, 工程 i の空き引き取りかんばんを $Q_i^{WK}(n)$, 工程 i の空き生産指示かんばんを $Q_i^{PK}(n)$, Q_i^D に残った需要 (工程 i での製品受注残) を $Q_i^D(n)$ とする。システムの観測後、製品の生産と需要の到着がイベントとして発生する (イベント3およびイベント4)。このときの工程 i の生産能力を $C_i(n)$, 実際に製品を生産した量を $P_i(n)$, 各工程に到着する需要を $D(n)$ とする。なお、 $C_i(n)$ は有限空間上の、 $D(n)$ は非負整数上の確率変数で、それぞれ次のように独立かつ同一な次の分布に従うものとする。

$$P(C_i(n) = j) = c_{i,j}, \quad j \in \mathcal{C}_i \equiv \{1, 2, \dots, C_i^{MAX}\}, \quad i \in \mathcal{F}, n \in \mathbb{N}$$

$$P(D(n) = j) = d_j, \quad j \in \mathcal{D} \equiv \{0, 1, 2, \dots\}, n \in \mathbb{N}$$

また、システムの初期状態として $Q^B(0) = 0$, $Q_i^D(0) = 0$, $Q_i^M(0) = S_i^M$, $Q_i^P(0) = S_i^P$, $Q_i^{WK}(0) = K_i^W - S_i^M =: S_i^{WK}$, $Q_i^{PK}(0) = K_i^P - S_i^P =: S_i^{PK}$ と定義する。また計算の便宜上、 $D(0) = 0$, $C_i(0) = 0$, $P_i(0) = 0$ と定義する。

3. システムの動き

期間 $n(n = 1, 2, \dots)$ でのシステムの動きを考える。まず期間 n の期首にイベント1が発生、すなわち工程 i が後工程 $\text{post}(i) = i'$ へと製品を発送する。このときに発送する量 $O_i^S(n)$ について考える。 $O_i^S(n)$ は需要量、空き引き取りかんばんの枚数、製品量の最小値である。期間 n のイベント1時点で工程 i に残っている顧客からの需要量すなわち期間 n の期首での Q_i^D に残ってる需要量は、期間 $n-1$ での受注残に期間 n に新規に到着した需要量を加えた量 $Q_i^D(n-1) + D(n-1)$ ($i \in \mathcal{P}$) となる。イ

ベント1時点で工程 $i \in \mathcal{P}_{-1}$ の引き取りかんばんポストにある空き引き取りかんばんの枚数は、期間 $n-1$ のシステムの状態観測点での空き引き取りかんばん枚数に、期間 $n-1$ での工程 i' の製品生産量を加えた量 $Q_i^{\text{WK}}(n-1) + P_{i'}(n-1)$ である。また、イベント1時点での工程 $i \in \mathcal{F}$ の製品置き場に置いてある製品量は期間 $n-1$ で置いてあった製品量に期間 $n-1$ で工程 i で製品を生産した量を加えた量 $Q_i^{\text{P}}(n-1) + P_i(n-1)$ である。イベント1で工程 i が発送する量はこれらの最小値であるので

$$O_i^{\text{S}}(n) = \min\{Q_i^{\text{D}}(n-1) + D(n-1), Q_i^{\text{WK}}(n-1) + P_{i'}(n-1), Q_i^{\text{P}}(n-1) + P_i(n-1)\}, \quad i \in \mathcal{P}, n \in \mathbb{N},$$

と表現できる。ただし、 $Q_1^{\text{WK}}(n) = \infty, Q_i^{\text{P}}(n) = \infty, i \in \mathcal{S}, n \in \mathbb{N}_0$ とする。特に、 $n=1$ のときは、 $Q_i^{\text{D}}(0) = D(0) = 0$ であることに気をつけると、 $O_i^{\text{S}}(1) = 0, (i \in \mathcal{P})$ となっていて、期間1では発送が行われないことを表している。

イベント2では各工程 i または顧客が製品を引き取る。このときに工程 i が j 番目の製品を受け取る量 $O_{i,j}^{\text{A}}(n)$ はイベント1でそれぞれ1つ前の工程が発送した量であるので

$$O_{i,j}^{\text{A}}(n) = O_{\text{pre}(i,j)}^{\text{S}}(n), \quad i \in \mathcal{F}, j \in [n_i], n \in \mathbb{N},$$

と表現できる。 O_i^{S} と同様に、 $n=1$ のときは $O_i^{\text{A}}(1) = 0, (i = 1, 2, \dots, k+1)$ となっていて、期間1では受取が行われないことを表している。

以下では、期間 n の状態観測点でのシステムの状態について考える。この時観測するシステムの状態は $Q_i^{\text{M}}(n), Q_j^{\text{P}}(n), Q_i^{\text{WK}}(n), Q_j^{\text{PK}}(n), Q_k^{\text{D}}(n) \ (i \in \mathcal{P}_{-1}, j \in \mathcal{F}, k \in \mathcal{P})$ である。これらの $5N+3M-2$ 個の状態が期間 $n-1$ の状態観測点から期間 n の状態観測点までのどれくらい変化したかを考える。まず、工程 $i \in \mathcal{P}_{-1}$ の部品在庫量 $Q_i^{\text{M}}(n)$ は、期間 $n-1$ のイベント3で製品生産量 $P_{i'}(n-1)$ だけ消費され、期間 n のイベント2で $O_{i', \text{postN}(i)}^{\text{A}}(n)$ ($\text{postN}(i)$ は工程 i の後工程内でのインデックス) だけ工程 i から部品が到着して増加するので、

$$Q_i^{\text{M}}(n) = Q_i^{\text{M}}(n-1) - P_{i'}(n-1) + O_{i', \text{postN}(i)}^{\text{A}}(n), \quad i \in \mathcal{F}, n \in \mathbb{N},$$

となる。工程 $i \in \mathcal{F}$ の製品在庫量 $Q_i^{\text{P}}(n)$ は、期間 $n-1$ のイベント3で製品生産量 $P_i(n-1)$ だけ生産され、期間 n のイベント1で $O_i^{\text{S}}(n)$ だけ工程 i' へ発送して減少するので、

$$Q_i^{\text{P}}(n) = Q_i^{\text{P}}(n-1) + P_i(n-1) - O_i^{\text{S}}(n), \quad i \in \mathcal{P}, n \in \mathbb{N},$$

となる。工程 $i \in \mathcal{P}_{-1}$ の引き取りかんばんポストにある空き引き取りかんばん枚数 $Q_i^{\text{WK}}(n)$ は、工程 i の引き取りかんばんの枚数 K_i^{W} から期間 n の状態観測点での部品在庫量 $Q_i^{\text{M}}(n)$ を引いたものになるので、

$$Q_i^{\text{WK}}(n) = K_i^{\text{W}} - Q_i^{\text{M}}(n), \quad i \in \mathcal{P}_{-1}, n \in \mathbb{N},$$

となる。工程 $i \in \mathcal{F}$ の生産指示かんばんポストにある空き生産指示かんばん枚数 $Q_i^{\text{PK}}(n)$ は、工程 i の生産指示かんばんの枚数 K_i^{P} から期間 n の状態観測点での製品在庫量 $Q_i^{\text{P}}(n)$ を引いたものになるので、

$$Q_i^{\text{PK}}(n) = K_i^{\text{P}} - Q_i^{\text{P}}(n), \quad i \in \mathcal{F}, n \in \mathbb{N},$$

となる。工程 $i \in \mathcal{P}$ の製品受注残 $Q_i^{\text{D}}(n)$ は期間 n のイベント4で新規に到着した需要量 $D(n-1)$ だけ追加され、期間 n のイベント1で製品発注に伴い製品発送量 $O_i^{\text{S}}(n)$ だけ需要が満たされることで減少するので、

$$Q_i^{\text{D}}(n) = Q_i^{\text{D}}(n-1) + D(n-1) - O_i^{\text{S}}(n), \quad i \in \mathcal{P}, n \in \mathbb{N},$$

となる。特に、 $n=1$ のときは、 $Q_i^{\text{D}}(1) = 0, Q_i^{\text{M}}(1) = S_i^{\text{M}}, Q_i^{\text{P}}(1) = S_i^{\text{P}}, Q_i^{\text{WK}}(1) = K_i^{\text{W}} - S_i^{\text{M}}, Q_i^{\text{PK}}(1) = K_i^{\text{P}} - S_i^{\text{P}}$ となり、初期状態から変化していない。

次にイベント3では各工程で製品が生産される。工程 $i \in \mathcal{F}$ での生産量 $P_i(n)$ は、工程 $j \in \text{pre}(i)$ の部品在庫量 $Q_j^M(n)$ を工程 i の空き生産指示かんばん枚数 $Q_i^{\text{PK}}(n)$ と工程 i の生産能力 $C_i(n)$ で制限した量となるので、

$$P_i(n) = \min\{C_i(n), Q_i^{\text{PK}}(n), \min\{Q_j^M(n) \mid j \in \text{pre}(i)\}\}, \quad i \in \mathcal{F}, n \in \mathbb{N}.$$

最後にイベント4では需要量 $D(n)$ が各工程の製品発注残 $Q_i^D(n)$, $i \in \mathcal{P}$ に加算される。

以上により、この節で述べた離散時間マルコフ連鎖をもちいたモデルは、 $5N + 3M - 2$ 個の状態

$$\{(Q_i^M(n), Q_j^P(n), Q_i^{\text{WK}}(n), Q_j^{\text{PK}}(n), Q_k^D(n)); n \geq 0, i \in \mathcal{P}_{-1}, j \in \mathcal{F}, k \in \mathcal{P}\}$$

で表現できることがわかる。

4. システムを表現する式

この節では3節で現れた式をまとめ、次節以降で用いる記号を定義する。

$$\begin{aligned} O_i^S(n) &= \min\{Q_i^D(n-1) + D(n-1), Q_i^{\text{WK}}(n-1) + P_{i'}(n-1), \\ &\quad Q_i^P(n-1) + P_i(n-1)\}, \quad i \in \mathcal{P}, n \in \mathbb{N}, \end{aligned}$$

$$O_{i,j}^A(n) = O_{\text{pre}(i,j)}^S(n), \quad i \in \mathcal{F}, j \in [n_i], n \in \mathbb{N},$$

特に

$$O_{i', \text{postN}(i)}^A(n) = O_i^S(n), \quad i \in \mathcal{F}, n \in \mathbb{N},$$

$$Q_i^M(n) = Q_i^M(n-1) - P_{i'}(n-1) + O_{i', \text{postN}(i)}^A(n), \quad i \in \mathcal{P}_{-1}, n \in \mathbb{N},$$

$$Q_i^P(n) = Q_i^P(n-1) + P_i(n-1) - O_i^S(n), \quad i \in \mathcal{P}, n \in \mathbb{N},$$

$$Q_i^{\text{WK}}(n) = K_i^W - Q_i^M(n), \quad i \in \mathcal{P}_{-1}, n \in \mathbb{N},$$

$$Q_i^{\text{PK}}(n) = K_i^P - Q_i^P(n), \quad i \in \mathcal{F}, n \in \mathbb{N},$$

$$Q_i^D(n) = Q_i^D(n-1) + D(n-1) - O_i^S(n), \quad i \in \mathcal{P}, n \in \mathbb{N},$$

$$P_i(n) = \min\{C_i(n), Q_i^{\text{PK}}(n), \min\{Q_j^M(n) \mid j \in \text{pre}(i)\}\}, \quad i \in \mathcal{F}, n \in \mathbb{N}.$$

ただし、システム上存在しない次の記号を計算の都合から下のように定義する。

$$\begin{aligned} Q_0^M(n) &= 0, \quad n \in \mathbb{N}_0, \\ Q_1^{\text{WK}}(n) &= \infty, \quad n \in \mathbb{N}_0, \\ P_0(n) &= \infty, \quad n \in \mathbb{N}, \\ Q_i^P(n) &= \infty, \quad i \in \mathcal{S}, n \in \mathbb{N}_0, \end{aligned}$$

また、計算のために次のように記号を定義する。

$$\begin{aligned} \bar{D}(n) &= \sum_{m=1}^n D(m), \quad n \in \mathbb{N}_0, \\ \bar{P}_i(n) &= \sum_{m=1}^n P_i(m), \quad i \in \mathcal{F}, n \in \mathbb{N}_0, \\ \bar{P}_i \bar{D}(n) &= \bar{P}_i(n) - \bar{D}(n), \quad i \in \mathcal{F}, n \in \mathbb{N}_0, \\ \bar{P}_i \bar{P}_j(n) &= \bar{P}_i(n) - \bar{P}_j(n), \quad i, j \in \mathcal{F}, n \in \mathbb{N}_0, \\ \bar{O}_i^S(n) &= \sum_{m=1}^n O_i^S(m), \quad i \in \mathcal{F} \cup \mathcal{S}, n \in \mathbb{N}_0, \end{aligned}$$

$$\begin{aligned}
\overline{Q}_i^D(n) &= S_i^D + \overline{D}(n-1), \quad i \in \mathcal{P}, n \in \mathbb{N}, \\
\overline{Q}_i^M(n) &= S_i^M + \overline{O}_i^S(n-1), \quad i \in \mathcal{P}, n \in \mathbb{N}, \\
\overline{Q}_i^P(n) &= S_i^P + \overline{P}_i(n-1), \quad i \in \mathcal{P}, n \in \mathbb{N}, \\
\overline{Q}_i^{\text{WK}}(n) &= S_i^{\text{WK}} + \overline{P}_{i'}(n-1), \quad i \in \mathcal{P}, n \in \mathbb{N}, \\
\overline{Q}_i^{\text{PK}}(n) &= S_i^{\text{PK}} + \overline{P}_i(n-1), \quad i \in \mathcal{F}, n \in \mathbb{N}.
\end{aligned}$$

ただし $\sum_{m=1}^0 \cdot = 0$ とする.

5. 主結果

本節では次の状態を考える.

$$\Pi(n) = Q_1^D(n), \Delta_i(n) = Q_i^D(n) - Q_{i'}^D(n), \quad i \in \mathcal{P}_{-1}$$

すると本研究の主張は次の定理である.

定理 1. 木構造を持つ多工程拡張かんばんシステムを表現する $\mathbf{Y}(n) = (\Pi(n), \Delta_i(n), Q_j^P(n), Q_k^M(n); i \in \mathcal{P}_{-1}, j \in \mathcal{P}, k \in \mathcal{F}, n \in \mathbb{N}_0)$ は M/G/1 型マルコフ連鎖に帰着できる.

$\mathbf{Y}(n)$ はマルコフ性を持ち, 主状態の 1 期間での最大減少幅が K_1^P であり, 主状態の最小値が 0 で上に非有界であることから次のことを示せば十分.

1. 背後状態の状態空間は有限である.
2. 主状態 $\Pi(n)$ が十分大きいとき背後状態の推移は $\Pi(n)$ に依存しない.

定理 2. 背後状態の状態空間は有限である.

明らかに $Q_i^P(n), Q_i^M(n)$ の状態空間は有限である. よって $\Delta_i(n)$ の状態空間が有限であることを示せば十分. そのためにいくつか補題を示す.

補題 1. $i \in \mathcal{P}_{-1}, n \in \mathbb{N}$ に対して次が成り立つ.

$$\overline{O}_i^S(n+1) \leq \overline{D}(n)$$

証明. $O_i^S(n+1)$ の定義に $Q_i^D(n), Q_i^{\text{WK}}(n), Q_i^P(n)$ の定義を次々に代入していくと,

$$\begin{aligned}
O_i^S(n+1) &= \min\{Q_i^D(n) + D(n), Q_i^{\text{WK}}(n) + P_{i'}(n), Q_i^P(n) + P_i(n)\} \\
&= \min\{Q_i^D(n-1) - O_i^S(n) + D(n-1) + D(n), \\
&\quad Q_i^{\text{WK}}(n-1) - O_i^S(n) + P_{i'}(n-1) + P_{i'}(n), \\
&\quad Q_i^P(n-1) - O_i^S(n) + P_i(n-1) + P_i(n)\} \\
&= \min\{Q_i^D(n-1) + D(n-1) + D(n), \\
&\quad Q_i^{\text{WK}}(n-1) + P_{i'}(n-1) + P_{i'}(n), \\
&\quad Q_i^P(n-1) + P_i(n-1) + P_i(n)\} - O_i^S(n) \\
&= \min\{Q_i^D(n-2) + D(n-2) + D(n-1) + D(n), \\
&\quad Q_i^{\text{WK}}(n-2) + P_{i'}(n-2) + P_{i'}(n-1) + P_{i'}(n), \\
&\quad Q_i^P(n-2) + P_i(n-2) + P_i(n-1) + P_i(n)\} - (O_i^S(n-1) + O_i^S(n)) \\
&= \cdots = \min\{Q_i^D(1) + \overline{D}(n), Q_i^{\text{WK}}(1) + \overline{P}_{i'}(n), Q_i^P(1) + \overline{P}_i(n)\} - \overline{O}_i^S(n) + O_i^S(1) \\
&= \min\{\overline{D}(n), S_i^{\text{WK}} + \overline{P}_{i'}(n), S_i^P + \overline{P}_i(n)\} - \overline{O}_i^S(n).
\end{aligned}$$

右辺の $\overline{O}_i^S(n)$ を左辺に移行すると,

$$\begin{aligned}
\overline{O}_i^S(n+1) &= \min\{\overline{D}(n), S_i^{\text{WK}} + \overline{P}_{i'}(n), S_i^P + \overline{P}_i(n)\} \\
&\leq \overline{D}(n) \quad \square
\end{aligned}$$

今示した補題 1 を用いて $Q_i^D(m)$ について評価する. $i \in \mathcal{P}_{-1}$ について

$$\begin{aligned} Q_i^D(m) &= Q_i^D(m-1) + D(m-1) - O_i^S(m), \\ Q_i^D(m) - Q_i^D(m-1) &= D(m-1) - O_i^S(m). \end{aligned}$$

両辺を $m = 1$ から n まで足し合わせると

$$\begin{aligned} Q_i^D(n) - Q_i^D(0) &= \overline{D}(n-1) - \overline{O}_i^S(n), \\ Q_i^D(n) &= \overline{D}(n-1) - \overline{O}_i^S(n). \end{aligned}$$

補題 1. から

$$Q_i^D(n) = \overline{D}(n-1) - \min\{\overline{D}(n-1), S_i^{\text{WK}} + \overline{P}_{i'}(n-1), S_i^{\text{P}} + \overline{P}_i(n-1)\}.$$

よって

$$\begin{aligned} \Delta_i(n) &= Q_i^D(n) - Q_{i'}^D(n) \\ &= \overline{D}(n-1) - \min\{\overline{D}(n-1), S_i^{\text{WK}} + \overline{P}_{i'}(n-1), S_i^{\text{P}} + \overline{P}_i(n-1)\} \\ &\quad - [\overline{D}(n-1) - \min\{\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}(n-1)\}] \\ &= -[\min\{\overline{D}(n-1), S_i^{\text{WK}} + \overline{P}_{i'}(n-1), S_i^{\text{P}} + \overline{P}_i(n-1)\} \\ &\quad + \max\{\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}(n-1)\}] \\ &= -\min \left\{ \begin{array}{l} \max[0, -S_{i'}^{\text{WK}} - \overline{P}_{i''}\overline{D}(n-1), -S_{i'}^{\text{P}} - \overline{P}_{i'}\overline{D}(n-1)], \\ S_i^{\text{WK}} + \max[\overline{P}_{i'}\overline{D}(n-1), -S_{i'}^{\text{WK}} - \overline{P}_{i''}\overline{P}_{i'}(n-1), -S_{i'}^{\text{P}}], \\ S_i^{\text{P}} + \max[\overline{P}_i\overline{D}(n-1), -S_{i'}^{\text{WK}} - \overline{P}_{i''}\overline{P}_i(n-1), -S_{i'}^{\text{P}} - \overline{P}_{i'}\overline{P}_i(n-1)] \end{array} \right\} \\ &= \max \left\{ \begin{array}{l} \min[0, S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{D}(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{D}(n-1)], \\ -S_i^{\text{WK}} + \min[-\overline{P}_{i'}\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{P}_{i'}(n-1), S_{i'}^{\text{P}}], \\ -S_i^{\text{P}} + \min[-\overline{P}_i\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{P}_i(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{P}_i(n-1)] \end{array} \right\}. \end{aligned}$$

ここで, 次を定義する.

$$\begin{aligned} X_i(n) &= \min[0, S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{D}(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{D}(n-1)], \\ Y_i(n) &= -S_i^{\text{WK}} + \min[-\overline{P}_{i'}\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{P}_{i'}(n-1), S_{i'}^{\text{P}}], \\ Z_i(n) &= -S_i^{\text{P}} + \min[-\overline{P}_i\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{P}_i(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{P}_i(n-1)] \\ &\quad i \in \mathcal{P}_{-1}, n \in \mathbb{N}. \end{aligned}$$

すると,

$$\Delta_i(n) = \max\{X_i(n), Y_i(n), Z_i(n)\}, \quad i \in \mathcal{P}_{-1}, n \in \mathbb{N}.$$

ここで, $\Delta_i(n)$ を評価するためにさらにいくつかの補題を示す.

補題 2. $i \in \mathcal{P}_{-1}, n \in \mathbb{N}$ に対して次が成り立つ.

$$\overline{P}_i\overline{D}(n-1) \in (-\infty, S_i^{\text{PK}}].$$

証明. Q_i^{PK} の定義から,

$$\begin{aligned} Q_i^{\text{PK}}(m+1) &= Q_i^{\text{PK}}(m) - P_i(m) + O_i^S(m+1), \\ P_i(m) &= Q_i^{\text{PK}}(m) - Q_i^{\text{PK}}(m+1) + O_i^S(m+1). \end{aligned} \tag{1}$$

(1) 式を m について 1 から n まで足し合わせると,

$$\begin{aligned}\overline{P}_i(n) &= Q_i^{\text{PK}}(1) - Q_i^{\text{PK}}(n+1) + \overline{O}_i^{\text{S}}(n+1) - O_i^{\text{S}}(1) \\ &= S_i^{\text{PK}} - Q_i^{\text{PK}}(n+1) + \overline{O}_i^{\text{S}}(n+1).\end{aligned}$$

よって補題 1. から,

$$\begin{aligned}\overline{P}_i\overline{D}(n-1) &= \overline{P}_i(n-1) - \overline{D}(n-1) \\ &\leq S_i^{\text{PK}} - Q_i^{\text{PK}}(n+1) + \overline{D}(n) - \overline{D}(n) \\ &\leq S_i^{\text{PK}}, \\ \therefore \overline{P}_i\overline{D}(n-1) &\in (-\infty, S_i^{\text{PK}}]. \quad \square\end{aligned}$$

補題 3. $i \in \mathcal{P}_{-1}$, $n \in \mathbb{N}$ に対して次が成り立つ.

$$\overline{P}_{i'}\overline{P}_i(n-1) \in [-\{S_i^{\text{PK}} + S_i^{\text{WK}}\}, S_i^{\text{P}} + S_i^{\text{M}}].$$

証明. $Q_i^{\text{P}}(n)$, $Q_i^{\text{M}}(n)$ の定義から,

$$\begin{aligned}Q_i^{\text{P}}(n+1) + Q_i^{\text{M}}(n+1) &= Q_i^{\text{P}}(n) + P_i(n) - O_i^{\text{S}}(n+1) + Q_i^{\text{M}}(n) - P_{i'}(n) + O_i^{\text{S}}(n+1) \\ &= P_i(n) - P_{i'}(n) + Q_i^{\text{P}}(n) + Q_i^{\text{M}}(n), \\ P_{i'}(n) - P_i(n) &= \{Q_i^{\text{P}}(n) + Q_i^{\text{M}}(n)\} - \{Q_i^{\text{P}}(n+1) + Q_i^{\text{M}}(n+1)\}, \\ \overline{P}_{i'}\overline{P}_i(n) &= \sum_{m=1}^n \{Q_i^{\text{P}}(m) + Q_i^{\text{M}}(m)\} \\ &= Q_i^{\text{P}}(1) + Q_i^{\text{M}}(1) - \{Q_i^{\text{P}}(n+1) + Q_i^{\text{M}}(n+1)\} \\ &= S_i^{\text{P}} - Q_i^{\text{P}}(n+1) + S_i^{\text{M}} - Q_i^{\text{M}}(n+1).\end{aligned}$$

ここで, $Q_i^{\text{P}}(n+1)$ の取りうる範囲を考えると,

$$\begin{aligned}Q_i^{\text{P}}(n+1) &\in [0, K_i^{\text{P}}] \\ &= [0, S_i^{\text{P}} + S_i^{\text{PK}}], \\ S_i^{\text{P}} - Q_i^{\text{P}}(n+1) &\in [-S_i^{\text{PK}}, S_i^{\text{P}}].\end{aligned}$$

同様に,

$$\begin{aligned}Q_i^{\text{M}}(n+1) &\in [0, K_i^{\text{M}}] \\ &= [0, S_i^{\text{M}} + S_i^{\text{WK}}], \\ S_i^{\text{M}} - Q_i^{\text{M}}(n+1) &\in [-S_i^{\text{WK}}, S_i^{\text{M}}].\end{aligned}$$

以上より,

$$\overline{P}_{i'}\overline{P}_i(n-1) \in [-\{S_i^{\text{PK}} + S_i^{\text{WK}}\}, S_i^{\text{P}} + S_i^{\text{M}}]. \quad \square$$

ここで, $\Delta_i(n)$ の各項について評価する.

補題 4. $i \in \mathcal{P}_{-1}$, $n \in \mathbb{N}$ に対して次が成り立つ.

$$X_i(n) \in (-\infty, 0].$$

証明. $X_i(n)$ の各項を評価する.

$$\begin{aligned}X_i(n) &= \min[0, S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{D}(n-1), S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{D}(n-1)], \\ S_{i'}^{\text{WK}} + \overline{P}_{i''}\overline{D}(n-1) &\in (-\infty, S_{i'}^{\text{WK}} + S_{i''}^{\text{PK}}], \\ S_{i'}^{\text{P}} + \overline{P}_{i'}\overline{D}(n-1) &\in (-\infty, S_{i'}^{\text{P}} + S_{i'}^{\text{PK}}], \\ \therefore X_i(n) &\in (-\infty, 0]. \quad \square\end{aligned}$$

補題 5. $i \in \mathcal{P}_{-1}$, $n \in \mathbb{N}$ に対して次が成り立つ.

$$Y_i(n) \in [-S_{i'}^{\text{PK}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} - S_i^{\text{WK}}].$$

証明. $Y_i(n)$ の各項を評価する.

$$\begin{aligned} Y_i(n) &= -S_i^{\text{WK}} + \min[-\overline{P_{i'}}\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P_{i''}}\overline{P_{i'}}(n-1), S_{i'}^{\text{P}}], \\ -\overline{P_{i'}}\overline{D}(n-1) &\in [-S_{i'}^{\text{PK}}, \infty), \\ S_{i'}^{\text{WK}} + \overline{P_{i''}}\overline{P_{i'}}(n-1) &\in [-S_{i'}^{\text{PK}}, S_{i'}^{\text{P}} + K_{i'}^{\text{W}}], \\ \therefore Y_i(n) &\in [-S_{i'}^{\text{PK}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} - S_i^{\text{WK}}]. \quad \square \end{aligned}$$

補題 6. $i \in \mathcal{P}_{-1}$, $n \in \mathbb{N}$ に対して次が成り立つ.

$$Z_i(n) \in [-S_{i'}^{\text{PK}} - K_i^{\text{P}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} + S_i^{\text{M}}].$$

証明. $Z_i(n)$ の各項を評価する.

$$\begin{aligned} Z_i(n) &= -S_i^{\text{P}} + \min[-\overline{P_i}\overline{D}(n-1), S_{i'}^{\text{WK}} + \overline{P_{i''}}\overline{P_i}(n-1), S_{i'}^{\text{P}} + \overline{P_{i'}}\overline{P_i}(n-1)], \\ -\overline{P_i}\overline{D}(n-1) &\in [-S_i^{\text{PK}}, \infty). \\ S_{i'}^{\text{WK}} + \overline{P_{i''}}\overline{P_i}(n-1) &\in [-S_{i'}^{\text{PK}} - S_i^{\text{PK}} - S_i^{\text{WK}}, K_{i'}^{\text{W}} + S_{i'}^{\text{P}} + S_i^{\text{P}} + S_i^{\text{M}}], \\ S_{i'}^{\text{P}} + \overline{P_{i'}}\overline{P_i}(n-1) &\in [S_{i'}^{\text{P}} - S_i^{\text{PK}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} + S_i^{\text{P}} + S_i^{\text{M}}], \\ \therefore Z_i(n) &\in [-S_{i'}^{\text{PK}} - K_i^{\text{P}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} + S_i^{\text{M}}]. \quad \square \end{aligned}$$

$\Delta_i(n)$ の各項について評価ができたので $\Delta_i(n)$ を評価する.

定理 3. $\Delta_i(n)$ は $i \in \mathcal{P}_{-1}$, $n \in \mathbb{N}$ に対して初期値にのみ依存した有限の範囲に収まる.

証明. 補題 4., 5., 6. より次が成り立つ.

$$\Delta_i(n) \in [-S_{i'}^{\text{PK}} - S_i^{\text{WK}}, S_{i'}^{\text{P}} + S_i^{\text{M}}] =: [\Delta_i^{\max}, \Delta_i^{\min}] \quad \square$$

次の定理を示す.

定理 4. $\Pi(n)$ が十分大きいとき背後状態の推移は $\Pi(n)$ に依存しない.

まず, $P_i(n+1)$ は定義から明らかに $\Pi(n)$ に依存しない. また, 次の補題が成り立つ.

補題 7. $\Pi(n) \geq \min(K_i^{\text{W}}, K_i^{\text{P}}) - \sum_{l \in \text{postAll}(i) \setminus \{1\}} \Delta_l^{\min}$, $i \in \mathcal{P}_{-1}$ を満たすとき, $O_i^{\text{S}}(n+1)$ の推移は $\Pi(n)$ に依存しない.

証明. $\text{postAll}(i)$ を工程 i の後工程から工程 1 までの工程全てのインデックスの集合とすると,

$$Q_1^{\text{D}}(n) + \sum_{l \in \text{postAll}(i) \setminus \{1\}} \Delta_l^{\min} = Q_i^{\text{D}}(n)$$

であるので, 仮定から

$$\begin{aligned} Q_1^{\text{D}}(n) + \sum_{l \in \text{postAll}(i) \setminus \{1\}} \Delta_l^{\min} &\geq \min(K_i^{\text{W}}, K_i^{\text{P}}) \\ Q_i^{\text{D}}(n) &\geq \min(K_i^{\text{W}}, K_i^{\text{P}}) \\ &= \min\{Q_i^{\text{WK}}(n) + Q_i^{\text{M}}(n), Q_i^{\text{P}}(n) + Q_i^{\text{PK}}(n)\} \\ &\geq \min\{Q_i^{\text{WK}}(n) + P_{i'}(n), Q_i^{\text{P}}(n) + P_i(n)\} \end{aligned}$$

であるので,

$$\begin{aligned} O_i^{\text{S}}(n+1) &= \min\{Q_i^{\text{D}}(n) + D(n), Q_i^{\text{WK}}(n) + P_{i'}(n), Q_i^{\text{P}}(n) + P_i(n)\} \\ &= \min\{Q_i^{\text{WK}}(n) + P_{i'}(n), Q_i^{\text{P}}(n) + P_i(n)\} \end{aligned}$$

よって示された. □

以上から定理 4. を示す.
証明 (定理 4.). 定義から

$$\begin{aligned}
& Q_i^M(n+1) \\
&= Q_i^M(n) - P_{i'}(n) + O_i^S(n+1), \\
& Q_i^P(n+1) \\
&= Q_i^P(n) + P_i(n) - O_i^S(n+1), \\
& \Delta_i(n+1) \\
&= \Delta_i(n) - O_i^S(n+1) + O_{i'}^S(n+1),
\end{aligned}$$

であるので非依存性が言えた. □

6. おわりに

本研究では, 木構造を持つ単一製品多工程拡張かんばんシステムについて離散時間待ち行列システムを用いることで, M/G/1 型のマルコフ連鎖に帰着できることを示した. 今後の課題として, 生産を行わない物流工程の考慮や, リードタイムが 0 でないモデルや多製品モデルへの拡張などが考えられる.

参考文献

- [1] 大野耐一: トヨタ生産方式—脱規模の経営をめざして—, ダイヤモンド社, 東京, 1978.
- [2] A.J. Clark and H. Scarf, "Optimal policies for multi-echelon inventory problem", Management Science, Vol. 6, pp.475–490, 1960.
- [3] M. L. Spearman, D. L. Woodruff and W. J. Hopp, "CONWIP: a pull alternative to kanban," International Journal of Production Research, vol. 28, no. 5, pp. 879–894, 1990.
- [4] A. M. Bonvik, C. E. Couch and S. B. Gershwin, "A comparison of production-line control mechanisms," International Journal of Production Research, vol. 35, no. 3, pp. 789–804, 1997.
- [5] Y. Dallery and G. Liberopoulos, "Extended kanban control system: Combining kanban and base stock," IIE Transactions, vol. 32, no. 4, pp. 369–386, 2000.
- [6] K. Ohno, "The optimal control of just-in-time-based production and distribution systems an performance comparisons with optimized pull systems," European Journal of Operational Research, vol. 213, no. 1, pp. 124–133, 2011.
- [7] O. Kimura and H. Terada, "Design and analysis of pull system, a method of multi-stage production control," International Journal of Production Research, vol. 19, no. 3, pp. 241–253, 1981.
- [8] N. Kirkavak and C. Dincer, "Performance evaluation models for single-item periodic pull production systems," Journal of the Operational Research Society, vol. 47, no. 2, pp. 239–250, 1996.
- [9] M. Iwase and K. Ohno, "The performance evaluation of a multi-stage JIT production system with stochastic demand and production capacities," European Journal of Operational Research, vol. 214, no. 2, pp. 216–222, 2011.
- [10] 影島るみ子, 井家敦, 佐久間大, "拡張かんばんシステムの推移構造について," 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集, pp. 162–163, 2015.
- [11] 牧本直樹: 待ち行列アルゴリズム—行列解析アプローチ—, 朝倉書店, 東京, 2001.

M/G/1 型マルコフ連鎖に対するレベル増分切断近似の漸近公式

大内 克久 増山 博之
京都大学大学院 情報学研究科

概要：本論文は、M/G/1 型マルコフ連鎖に対するレベル増分切断近似の収束速度について議論する。M/G/1 型マルコフ連鎖の定常分布ベクトルは、M/G/1 パラダイムによって計算可能である。しかし、M/G/1 パラダイムの実装には、レベル増分切断、すなわち、一推移あたりのレベル増分を切断する必要がある。そして、レベル増分切断の結果得られる定常分布は真の定常分布の近似となる。本論文では、そうした切断誤差に着目し、切断パラメータを無限に大きくしたときの漸近特性を調べる。まず、レベル増分分布の平衡分布が長裾的だと仮定し、その裾減衰と、レベル増分切断によって生じる定常分布ベクトルの差分の減衰が漸近的に等価であることを示す。さらに、レベル増分分布の平衡分布が劣指数的である場合には、定常分布ベクトルの差分の減衰と真の定常分布の裾減衰が漸近的に等価であることを示す。なお、レベル増分分布が軽裾的である場合についても、いくつかの漸近結果を与える。

1. 序論

M/G/1 型マルコフ連鎖は、セミ・マルコフ型待ち行列の解析を支える構造化マルコフ連鎖の一つである。例えば、マルコフ型集団到着過程 (BMAP: Batch Markovian Arrival Process) [5] と一般的なサービス時間分布を入力にもつ先着順単一サーバ待ち行列や、位相型サービス時間分布、あるいは、固定長サービスと BMAP を入力にもつ先着順複数サーバ待ち行列などの解析は、通常、M/G/1 型マルコフ連鎖を介して行われる。

M/G/1 型マルコフ連鎖の定常分布ベクトルは、Ramaswami の再帰式 [14, 15] を土台とする M/G/1 パラダイム [13] によって計算可能である。しかし、M/G/1 パラダイムを実装するには、推移を特徴づけるブロック行列の無限列を切断しなければならない。この切断は、一推移あたりのレベル増分の切断と等価であり、これを「レベル増分切断」とよぶ。なお、レベル増分切断によって得られるマルコフ連鎖も、M/G/1 型となる。

ところで、レベル増分切断によって得られる定常分布は、元の M/G/1 型マルコフ連鎖の定常分布に対する近似である。本論文の目的は、レベル増分切断による定常分布の誤差に着目し、切断パラメータを無限に大きくしたときの漸近特性を調べることである。以下、関連研究について述べる。

Miyazawa ら [12] は、準出生死滅過程 (QBD: Quasi-Birth-and-Death process) の到達レベルに上限を設けた有限レベル QBD を考えている。Miyazawa ら [12] は、有限レベル QBD の上限レベルを無限に大きくしたときの定常分布の漸近公式を導出している。ちなみに、無限レベル (すなわち、通常の) QBD は、特別な M/G/1 型マルコフ連鎖であり、1 回の推移で高々 1 しかレベルは変化しない。さらに、Kim ら [4] は、Miyazawa ら [12] の結果を、GI/M/1 型マルコフ連鎖に拡張している。

有限レベル M/G/1 型マルコフ連鎖を対象とした先行研究もある。ちなみに、有限レベル M/G/1 型マルコフ連鎖とは、到達レベルに上限のある M/G/1 型マルコフ連鎖を指す。Baicocchi [2] は、マルコフ型到着過程 (集団到着のない BMAP) と一般サービス時間分布をもつ単一サーバ有限待ち行列の呼損率の漸近解析を行っている。この待ち行列の系内客数過程は、客の退去時刻で観測すると、有限レベル M/G/1 型マルコフ連鎖となる。Masuyama ら [11] は、有限レベル M/G/1 型マルコフ連鎖とそれに対応する無限レベル (すなわち、通常の) M/G/1 型マルコフ連鎖、これら両者の定常分布ベクトルの差分に対して、上限レベルを無限に大きくしたときの漸近公式を導いている。具体的には、レベル増分分布の平衡分布が長裾的である場合、差分ベクトルがレベル増分分布の平衡分布の裾減衰と同等の速度で減衰することを示している。さらに、レベル増分分布の平衡分布が劣指数的である場合には、同差分ベクトルの減衰が無限レベル M/G/1 型マルコフ連鎖の定常分布ベクトルの裾減衰と同等であることを明らかにしている。

その他の関連研究として、切断近似による定常分布の誤差上界を対象とするものがある。上で言及した有限レベル M/G/1 型マルコフ連鎖は、対応する無限レベル M/G/1 型マルコフ連鎖の最終列ブロック増大切断とみなすことができる。この最終列ブロック増大切断、および、より一般的な切断法によって得られる定常分布に対して、計算可能な誤差上界式がいくつか示されている [6, 7, 9, 10]。

さて、本論文では、Masuyama ら [11] が示した漸近公式と類似する結果が、M/G/1 型マルコフ連鎖のレベル増分切断に対しても成立することを示す。さらに、レベル増分分布の平衡分布が幾何的となる場合についても、同様の漸近公式を導く。

本論文の構成は以下のとおりである。第 2 節で M/G/1 型マルコフ連鎖とそのレベル増分切断について記述する。第 3 節および第 4 節では、切断誤差が劣幾何的に減衰する場合と幾何的に減衰する場合について、それぞれ漸近公式を示す。

2. M/G/1 型マルコフ連鎖

本節では、まず、M/G/1 型マルコフ連鎖の定常分布に関する Ramaswami の再帰式を紹介する。つづいて、Ramaswami の再帰式を実装する際に必要となるレベル増分切断を説明する。なお、以下では、整数全体の集合を表す $\mathbb{Z}$ に加え、次のような $\mathbb{Z}$ の部分集合を表すいくつかの記号を用いる。

$$\begin{aligned}\mathbb{Z}_+ &= \{n \in \mathbb{Z} : n \geq 0\}, & \mathbb{N} &= \{n \in \mathbb{Z} : n \geq 1\}, \\ \mathbb{Z}_{\geq k} &= \{n \in \mathbb{Z} : n \geq k\}, & \mathbb{Z}_{[k, \ell]} &= \{n \in \mathbb{Z} : k \leq n \leq \ell\}.\end{aligned}$$

ここで、 k, ℓ は実数とする。

2.1. Ramaswami の再帰式

2 変数確率過程 $\{(X_n, J_n); n \in \mathbb{Z}_+\}$ を状態空間 $\mathbb{S} := \bigcup_{k=0}^{\infty} \{k\} \times \mathbb{M}_{k \wedge 1}$ 上のマルコフ連鎖とする。ただし、 $x \wedge y = \min(x, y)$, $\mathbb{M}_0 = \{1, 2, \dots, M_0\} \subset \mathbb{N}$, $\mathbb{M}_1 = \{1, 2, \dots, M_1\} \subset \mathbb{N}$ である。さらに、状態部分集合 $\mathbb{L}_k$ ($k \in \mathbb{Z}_+$) を、

$$\mathbb{L}_k = k \times \mathbb{M}_{k \wedge 1}, \quad k \in \mathbb{Z}_+,$$

と定義し、これをレベル k とよぶ。ここで、マルコフ連鎖 $\{(X_n, J_n)\}$ の推移確率行列 $\mathbf{P}$ は、次のようなブロック構造をもつと仮定する。

$$\mathbf{P} = \begin{matrix} & \mathbb{L}_0 & \mathbb{L}_1 & \mathbb{L}_2 & \mathbb{L}_3 & \cdots \\ \begin{matrix} \mathbb{L}_0 \\ \mathbb{L}_1 \\ \mathbb{L}_2 \\ \mathbb{L}_3 \\ \vdots \end{matrix} & \begin{pmatrix} \mathbf{B}(0) & \mathbf{B}(1) & \mathbf{B}(2) & \mathbf{B}(3) & \cdots \\ \mathbf{B}(-1) & \mathbf{A}(0) & \mathbf{A}(1) & \mathbf{A}(2) & \cdots \\ \mathbf{O} & \mathbf{A}(-1) & \mathbf{A}(0) & \mathbf{A}(1) & \cdots \\ \mathbf{O} & \mathbf{O} & \mathbf{A}(-1) & \mathbf{A}(0) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{matrix}. \quad (2.1)$$

このとき、 $\{(X_n, J_n)\}$ は、M/G/1 型マルコフ連鎖とよばれる [13]。なお、 $\mathbf{P}$ は確率行列であるので、行列 $\mathbf{A} := \sum_{k=-1}^{\infty} \mathbf{A}(k)$ も確率行列である。すなわち、

$$\mathbf{A}\mathbf{e} = \sum_{k=-1}^{\infty} \mathbf{A}(k)\mathbf{e} = \mathbf{e}, \quad (2.2)$$

が成り立つ。同様に、

$$\mathbf{B}(0)\mathbf{e} + \sum_{k=1}^{\infty} \mathbf{B}(k)\mathbf{e} = \mathbf{e}, \quad (2.3)$$

$$\mathbf{B}(-1)\mathbf{e} + \sum_{k=0}^{\infty} \mathbf{A}(k)\mathbf{e} = \mathbf{e}, \quad (2.4)$$

が成り立つ。なお、後の議論ため、ここで、

$$\begin{aligned}\overline{\mathbf{A}}(k) &= \sum_{\ell=k+1}^{\infty} \mathbf{A}(\ell), & k &\in \mathbb{Z}_{\geq -2}, \\ \overline{\mathbf{B}}(k) &= \sum_{\ell=k+1}^{\infty} \mathbf{B}(\ell), & k &\in \mathbb{Z}_{\geq -1},\end{aligned}$$

と定義しておく. また, 式 (2.1) のように, 行列をレベルごとに分割したとき, レベル k に属する行とレベル ℓ に属する列からなる部分行列を, (k, ℓ) -ブロックとよぶ.

以下では, 推移確率行列 $\mathbf{P}$ の定常分布を考えるため, 次の仮定をおく.

仮定 2.1 (i) 推移確率行列 $\mathbf{P}$ は既約かつ非周期である. (ii) $\mathbf{A}$ は既約である. (iii) $\sum_{k=1}^{\infty} k \mathbf{B}(k) \mathbf{e} < \infty$ である. (iv) $\sigma := \varpi \sum_{k=-1}^{\infty} k \mathbf{A}(k) \mathbf{e} < 0$ を満たす. ただし, ϖ を $\mathbf{A}$ の唯一の定常分布ベクトルとする.

仮定 2.1 のもとで, 推移確率行列 $\mathbf{P}$, すなわち, M/G/1 型マルコフ連鎖 $\{(X_n, J_n)\}$ は, 唯一の定常分布ベクトル $\boldsymbol{\pi} = (\pi(k, i))_{(k, i) \in \mathbb{S}}$ をもつ [1, Chapter XI, Proposition 3.1]. ここで, $\boldsymbol{\pi}(k) = (\pi(k, i))_{i \in \mathbb{M}_{k \wedge 1}}$ ($k \in \mathbb{Z}_+$) とおくと, 定常分布ベクトル $\boldsymbol{\pi}$ を, 次のようにレベル単位で分割することができる.

$$\begin{array}{ccc} \mathbb{L}_0 & \mathbb{L}_1 & \mathbb{L}_2 \\ \boldsymbol{\pi} = (\boldsymbol{\pi}(0), \boldsymbol{\pi}(1), \boldsymbol{\pi}(2), \dots). \end{array}$$

注意 2.2 $\mathbf{P}$ の非周期性は, 定常分布ベクトル $\boldsymbol{\pi}$ の存在条件には直接関係はないが, 補題 2.5 の証明で用いられる.

ベクトル列 $\{\boldsymbol{\pi}(k); k \in \mathbb{Z}_+\}$ については, Ramaswami [14] によって示された再帰式 (Ramaswami の再帰式) が知られている. その再帰式を記述するため, いくつかの記号を定義する. まずは, 最も重要な G 行列の導入から始める. 任意の $k \in \mathbb{Z}_+$ に対して,

$$T_{\leq k} = \inf\{n \in \mathbb{N} : X_n \leq k\},$$

とする. また, 集合 $\mathbb{L}_{\geq k}$ と $\mathbb{L}_{\leq k}$ をそれぞれ,

$$\mathbb{L}_{\geq k} = \bigcup_{\ell=k}^{\infty} \mathbb{L}_{\ell}, \quad \mathbb{L}_{\leq k} = \bigcup_{\ell=0}^k \mathbb{L}_{\ell}, \quad k \in \mathbb{Z}_+,$$

とおく. これらを用いて, 行列 $\mathbf{G} := (G_{i,j})_{i,j \in \mathbb{M}_1}$ を次のように定める.

$$G_{i,j} = \mathbf{P}((X_{T_{\leq \ell}}, J_{T_{\leq \ell}}) = (\ell, j) \mid (X_0, J_0) = (\ell + 1, i) \in \mathbb{L}_{\geq 2}), \quad i, j \in \mathbb{M}_1.$$

仮定 2.1 の (ii) と (iv) より, $\mathbf{G}$ は確率行列であり [13, Theorem 2.3.1], その定常分布ベクトルを $\mathbf{g}$ とおく. なお, 行列 $\mathbf{G}$ は次の行列方程式の最小非負解であることが示されている [13, Theorem 2.2.2].

$$\mathbf{G} = \sum_{m=0}^{\infty} \mathbf{A}(m-1) \mathbf{G}^m. \quad (2.5)$$

つづいて, 行列 $\mathbf{G}$ を用いて, いくつか行列を定義する. 行列 $\mathbf{K} := (K_{i,j})_{i,j \in \mathbb{M}_0}$ を,

$$\mathbf{K} = \mathbf{B}(0) + \sum_{m=1}^{\infty} \mathbf{B}(m) \mathbf{G}^{m-1} (\mathbf{I} - \boldsymbol{\Phi}(0))^{-1} \mathbf{B}(-1), \quad (2.6)$$

と定義する. 仮定 2.1 の下で, $\mathbf{K}$ は確率行列となり ([15, Lemma 2.1]), その定常分布ベクトルを $\boldsymbol{\kappa}$ と表記する. さらに, $\boldsymbol{\Phi}(0) := (\Phi_{i,j}(0))_{i,j \in \mathbb{M}_1}$ を,

$$\boldsymbol{\Phi}(0) = \sum_{m=0}^{\infty} \mathbf{A}(m) \mathbf{G}^m, \quad (2.7)$$

と定め, これと行列 $\mathbf{G}$ を用いて, 行列 $\mathbf{R}(k) := (R_{i,j}(k))_{(i,j) \in (\mathbb{M}_1)^2}$ と $\mathbf{R}_0(k) := (R_{0,i,j}(k))_{(i,j) \in \mathbb{M}_0 \times \mathbb{M}_1}$ ($k \in \mathbb{N}$) を, それぞれ次のように定義する.

$$\mathbf{R}(k) = \sum_{m=0}^{\infty} \mathbf{A}(k+m) \mathbf{G}^m (\mathbf{I} - \boldsymbol{\Phi}(0))^{-1}, \quad k \in \mathbb{N}, \quad (2.8)$$

$$\mathbf{R}_0(k) = \sum_{m=0}^{\infty} \mathbf{B}(k+m) \mathbf{G}^m (\mathbf{I} - \boldsymbol{\Phi}(0))^{-1}, \quad k \in \mathbb{N}. \quad (2.9)$$

なお、上で定義した行列の各成分は以下のような意味をもつ。

$$\begin{aligned} K_{i,j} &= \mathbf{P}((X_{T_{\leq 0}}, J_{T_{\leq 0}}) = (0, j) \mid (X_0, J_0) = (0, i)), \\ \Phi_{i,j}(0) &= \mathbf{P}((X_{T_{\leq \ell}}, J_{T_{\leq \ell}}) = (\ell, j) \mid (X_0, J_0) = (\ell, i) \in \mathbb{L}_{\geq 1}), \\ R_{i,j}(k) &= \mathbf{E} \left[\sum_{n=1}^{T_{\leq \ell+k-1}} \mathbb{1}\{(X_n, J_n) = (\ell+k, j)\} \mid (X_0, J_0) = (\ell, i) \in \mathbb{L}_{\ell} \right], \\ R_{0,i,j}(k) &= \mathbf{E} \left[\sum_{n=1}^{T_{\leq k-1}} \mathbb{1}\{(X_n, J_n) = (k, j)\} \mid (X_0, J_0) = (0, i) \in \mathbb{L}_0 \right]. \end{aligned}$$

以上の準備のもと、ベクトル列 $\{\pi(k); k \in \mathbb{Z}_+\}$ に関する Ramaswami の再帰式を与える。

命題 2.3 ([14], [15, Theorem 3.1]) 任意の $k \in \mathbb{N}$ に対して、

$$\pi(k) = \pi(0)R_0(k) + \sum_{\ell=1}^{k-1} \pi(\ell)R(k-\ell), \quad k \in \mathbb{N},$$

が成り立つ。なお、 $\pi(0)$ は次式で与えられる。

$$\pi(0) = \kappa \left[1 + \frac{\kappa}{-\sigma} \left\{ \beta_B + \left(\sum_{m=1}^{\infty} B(m)(I - G^m) \right) (I - A + e\varpi)^{-1} \beta_A \right\} \right]^{-1}. \quad (2.10)$$

ただし、

$$\beta_A = \sum_{m=-1}^{\infty} m A(m) e, \quad \beta_B = \sum_{m=1}^{\infty} m B(m) e,$$

である。

2.2. Ramaswami の再帰式の実装とレベル増分切断

M/G/1 型マルコフ連鎖の定常分布ベクトルは、上の命題で示された Ramaswami の再帰式を基に計算される。しかし、Ramaswami の再帰式の実装には、いくつか問題がある。そのなかで最も本質的なものは、無限和の取り扱いである。

Ramaswami の再帰式を構成するには、行列 G が必要である。通常、行列 G の計算は、以下の再帰式、あるいはそれを改良した再帰式で生成される行列の列 $\{G(n); n \in \mathbb{Z}_+\}$ の極限を求めることで実現される [3]。

$$G(0) = O, \quad G(n) = \sum_{m=0}^{\infty} A(m-1)[G(n-1)]^m, \quad n \in \mathbb{N}. \quad (2.11)$$

しかし、上の再帰式には無限列 $\{A(k); k \in \mathbb{Z}_{\geq -1}\}$ に関する無限和が含まれる。また、行列 K の定義式 (2.6) にも、無限列 $\{B(k); k \in \mathbb{Z}_+\}$ に関する無限和が含まれている。そこで、無限列 $\{A(k)\}$ と $\{B(k)\}$ をそれぞれ有限列に置き換える近似手法がよく用いられる。

この近似手法では、まず、切断パラメータ N ($N \in \mathbb{N}$) を定め、行列 P の定義式 (2.1) において、次のような置き換えを行って得られる確率行列 $P^{(N)}$ を考える。

$$\begin{aligned} A(k) &= \begin{cases} \tilde{A}(N-1), & k = N, \\ O, & k \in \mathbb{Z}_{\geq N+1}, \end{cases} \\ B(k) &= \begin{cases} \tilde{B}(N-1), & k = N, \\ O, & k \in \mathbb{Z}_{\geq N+1}. \end{cases} \end{aligned}$$

ただし, $\tilde{\mathbf{A}}(k)$ ($k \in \mathbb{Z}_{\geq -2}$) と $\tilde{\mathbf{B}}(k)$ ($k \in \mathbb{Z}_{\geq -1}$) は, 次式を満たす劣確率的な行列である.

$$\begin{aligned}\tilde{\mathbf{A}}(k)\mathbf{e} &= \overline{\mathbf{A}}(k)\mathbf{e} = \sum_{\ell=k+1}^{\infty} \mathbf{A}(\ell)\mathbf{e}, & k \in \mathbb{Z}_{\geq -2}, \\ \tilde{\mathbf{B}}(k)\mathbf{e} &= \overline{\mathbf{B}}(k)\mathbf{e} = \sum_{\ell=k+1}^{\infty} \mathbf{B}(\ell)\mathbf{e}, & k \in \mathbb{Z}_{\geq -1}.\end{aligned}$$

確率行列 $\mathbf{P}^{(N)}$ の定義を正確に記述すると, その (k, ℓ) -ブロック $\mathbf{P}_{k, \ell}^{(N)}$ ($k, \ell \in \mathbb{Z}_+$) は次式で与えられる.

$$\mathbf{P}_{k, \ell}^{(N)} = \begin{cases} \mathbf{B}(\ell), & k = 0, \quad 0 \leq \ell \leq N-1, \\ \tilde{\mathbf{B}}(N-1), & k = 0, \quad \ell = N, \\ \mathbf{B}(-1), & k = 1, \quad \ell = 0, \\ \mathbf{A}(-1), & k \geq 2, \quad \ell - k = -1, \\ \mathbf{A}(\ell - k), & k \geq 1, \quad 0 \leq \ell - k \leq N, \\ \tilde{\mathbf{A}}(N-1), & k \geq 1, \quad \ell - k = N, \\ \mathbf{O}, & \text{その他.} \end{cases} \quad (2.12)$$

これより, $\mathbf{P}^{(N)}$ は, 一推移あたりのレベル増分が高々 N に制限された **M/G/1** 型確率行列であり,

$$\lim_{N \rightarrow \infty} \mathbf{P}^{(N)} = \mathbf{P}, \quad (2.13)$$

を満たすことがわかる. したがって, 本論文では, $\mathbf{P}^{(N)}$ を $\mathbf{P}$ に対するレベル増分切断近似とよぶ.

レベル増分切断近似 $\mathbf{P}^{(N)}$ は, 仮定 2.1 のもとで, 定常分布ベクトル $\boldsymbol{\pi}^{(N)} := (\pi^{(N)}(k, i))_{(k, i) \in \mathbb{F}}$ をもつ. さらに, $\boldsymbol{\pi}^{(N)}(k) = (\pi^{(N)}(k, i))_{(k, i) \in \mathbb{L}_k}$ ($k \in \mathbb{Z}_+$) と定義すると, 部分ベクトル列 $\{\boldsymbol{\pi}^{(N)}(k); k \in \mathbb{Z}_+\}$ に対して Ramaswami の再帰式が成り立ち, それは計算機に実装することができる.

まず, **M/G/1** 型確率行列 $\mathbf{P}^{(N)}$ に対する G 行列は, 以下の再帰式で生成される行列の列 $\{\mathbf{G}^{(N)}(n); n \in \mathbb{Z}_+\}$ の極限である.

$$\begin{aligned}\mathbf{G}^{(N)}(0) &= \mathbf{O}, \\ \mathbf{G}^{(N)}(n) &= \sum_{m=0}^N \mathbf{A}(m-1)[\mathbf{G}^{(N)}(n-1)]^m \\ &\quad + \tilde{\mathbf{A}}(N-1)[\mathbf{G}^{(N)}(n-1)]^{N+1}, \quad n \in \mathbb{N}.\end{aligned}$$

式 (2.11) における無限和が, この再帰式では有限和に置き換わっていることに注意する.

くわえて, $\mathbf{P}^{(N)}$ に対しても, $\mathbf{K}$, $\boldsymbol{\Phi}$, $\mathbf{R}(k)$, $\mathbf{R}_0(k)$ ($k \in \mathbb{N}$) に対応する行列 $\mathbf{K}^{(N)}$, $\boldsymbol{\Phi}^{(N)}$, $\mathbf{R}^{(N)}(k)$, $\mathbf{R}_0^{(N)}(k)$ ($k \in \mathbb{N}$) を定義することができ, それらは次式で与えられる.

$$\begin{aligned}\mathbf{K}^{(N)} &= \mathbf{B}(0) + \left[\sum_{m=1}^{N-1} \mathbf{B}(m)[\mathbf{G}^{(N)}]^{m-1} + \tilde{\mathbf{B}}(N-1)[\mathbf{G}^{(N)}]^{N-1} \right] (\mathbf{I} - \boldsymbol{\Phi}(0)^{(N)})^{-1} \mathbf{B}(-1), \\ \boldsymbol{\Phi}^{(N)}(0) &= \sum_{m=0}^{N-1} \mathbf{A}(m)[\mathbf{G}^{(N)}]^m + \tilde{\mathbf{A}}(N-1)[\mathbf{G}^{(N)}]^N, \\ \mathbf{R}^{(N)}(k) &= \left[\sum_{m=0}^{N-k-1} \mathbf{A}(k+m)[\mathbf{G}^{(N)}]^m + \tilde{\mathbf{A}}(N-1)[\mathbf{G}^{(N)}]^{N-k} \right] (\mathbf{I} - \boldsymbol{\Phi}(0)^{(N)})^{-1}, \quad k \in \mathbb{N}, \\ \mathbf{R}_0^{(N)}(k) &= \left[\sum_{m=0}^{N-k-1} \mathbf{B}(k+m)[\mathbf{G}^{(N)}]^m + \tilde{\mathbf{B}}(N-1)[\mathbf{G}^{(N)}]^{N-k} \right] (\mathbf{I} - \boldsymbol{\Phi}(0)^{(N)})^{-1}, \quad k \in \mathbb{N}.\end{aligned}$$

明らかに, これらの行列は数値的に求めることができる.

以上のことから, 部分ベクトル列 $\{\pi^{(N)}(k)\}$ に対する Ramaswami の再帰式の構成要素はすべて数値計算が可能である. なお, 当該再帰式は, これまでの議論と命題 2.3 から直ちに得ることができ, かつ, 後の議論に直接関係がないので, 割愛する.

さて, 式 (2.13) から期待されるように, 適当な条件のもとで切断パラメータ N を無限に大きくすると, レベル増分切断近似 $\mathbf{P}^{(N)}$ の定常分布ベクトル $\pi^{(N)}$ は, 元の定常分布ベクトル π に収束する. この結果を述べるため, 仮定 2.1 に加えて次の仮定をおく.

仮定 2.4

$$\sum_{k=1}^{\infty} k^2 \mathbf{A}(k) < \infty, \quad \sum_{k=1}^{\infty} k^2 \mathbf{B}(k) < \infty.$$

補題 2.5 仮定 2.1 および 2.4 のもとで,

$$\lim_{N \rightarrow \infty} \pi^{(N)} = \pi, \quad (2.14)$$

が成り立つ.

3. 劣幾何的漸近公式

この節では, $\pi^{(N)}$ が, 劣幾何的な速度で π に収束する場合を考える. その議論に必要な離散型分布をいくつか導入する. なお, 以下では, 任意の分布関数 F に対して, その補分布関数を $\bar{F}$ と表記する.

定義 3.1 すべての $k \in \mathbb{Z}_+$ に対して, $\bar{F}(k) > 0$ であり, かつ,

$$\lim_{k \rightarrow \infty} \frac{\bar{F}(k+n)}{\bar{F}(k)} = 1, \quad \forall n \in \mathbb{Z}_+,$$

であるとき, 分布関数 F を裾の長い分布という.

定義 3.2 すべての $k \in \mathbb{Z}_+$ に対して, $\bar{F}(k) > 0$ であり, かつ,

$$\lim_{k \rightarrow \infty} \frac{\bar{F}^{*2}(k)}{\bar{F}(k)} = 2,$$

であるとき, 分布関数 $F(x)$ は劣指数分布であるという. ただし, $F^{*1} = F$ であり, $n = 2, 3, \dots$ に対して, F^{*n} は n 重畳み込みを表す. つまり,

$$F^{*n}(k) = \sum_{\ell=0}^k F^{*(n-1)}(k-\ell)F(\ell), \quad k \in \mathbb{Z}_+,$$

である.

さらに, 裾の長い分布の族を $\mathcal{L}$, 劣指数分布の族を $\mathcal{S}$ と表す.

注意 3.1 劣指数分布族 $\mathcal{S}$ は $\mathcal{L}$ の部分族である.

ついで, $\widetilde{\mathbf{A}}(k)$ ($k \in \mathbb{Z}_{\geq -3}$) および $\widetilde{\mathbf{B}}(k)$ ($k \in \mathbb{Z}_{\geq -2}$) を

$$\begin{aligned} \widetilde{\mathbf{A}}(k) &= \sum_{\ell=k+1}^{\infty} \widetilde{\mathbf{A}}(\ell), & k \in \mathbb{Z}_{\geq -3}, \\ \widetilde{\mathbf{B}}(k) &= \sum_{\ell=k+1}^{\infty} \widetilde{\mathbf{B}}(\ell), & k \in \mathbb{Z}_{\geq -2}, \end{aligned}$$

と定義し, 以下の仮定をおく.

仮定 3.2 以下を満たす分布関数 $F \in \mathcal{L}$ と定ベクトル c_A, c_B が存在する.

$$\lim_{N \rightarrow \infty} \frac{\overline{\mathbf{A}}(N)e}{\overline{F}(N)} = c_A, \quad \lim_{N \rightarrow \infty} \frac{\overline{\mathbf{B}}(N)e}{\overline{F}(N)} = c_B.$$

ただし, c_A, c_B はそれぞれ M 次, M_0 次の列ベクトルであり, 少なくともどちらか一方は非ゼロベクトルである.

ベクトル $\overline{\pi}(k)$ ($k \in \mathbb{Z}_+$) を,

$$\overline{\pi}(k) = \sum_{\ell=k+1}^{\infty} \pi(\ell), \quad k \in \mathbb{Z}_+,$$

と定義すると, 補題 2.5 を用いて以下の定理を示すことができる.

定理 3.3 仮定 2.1, 2.4 および 3.2 のもとで,

$$\lim_{N \rightarrow \infty} \frac{\pi^{(N)} - \pi}{\overline{F}(N)} = \frac{\pi(0)c_B + \overline{\pi}(0)c_A}{-\sigma} \pi, \quad (3.1)$$

が成り立つ.

次に, 定理 3.3 を用いて, 元の M/G/1 型マルコフ連鎖の定常分布の裾 $\overline{\pi}(N)$ と $\pi^{(N)} - \pi$ の収束速度に関する漸近公式を導出する. ここで, その導出に用いる漸近公式を紹介する.

命題 3.4 ([8, Theorem 3.1]) 仮定 2.1 および 3.2 のもとで, $F \in \mathcal{S}$ であるならば,

$$\lim_{N \rightarrow \infty} \frac{\overline{\pi}(N)}{\overline{F}(N)} = \frac{\pi(0)c_B + \overline{\pi}(0)c_A}{-\sigma} \varpi,$$

が成り立つ.

定理 3.3 と命題 3.4 より, 次の系が得られる.

系 3.5 仮定 2.1, 2.4 および, 3.2 のもとで, $F \in \mathcal{S}$ であるならば,

$$\lim_{N \rightarrow \infty} \frac{\pi^{(N)} - \pi}{\overline{\pi}(N)e} = \pi,$$

が成り立つ.

4. 幾何的漸近公式

この節では, $\pi^{(N)}$ が, 幾何的な速度で π に収束する場合を考える. まず, $\hat{\mathbf{A}}(z), \hat{\mathbf{B}}(z)$ をそれぞれ $\{\mathbf{A}(k); k \in \mathbb{Z}_{\geq -1}\}, \{\mathbf{B}(k); k \in \mathbb{N}\}$ の確率母関数とする. すなわち,

$$\hat{\mathbf{A}}(z) = \sum_{k=-1}^{\infty} z^k \mathbf{A}(k), \quad \hat{\mathbf{B}}(z) = \sum_{k=1}^{\infty} z^k \mathbf{B}(k),$$

とする. さらに, 確率母関数 $\hat{\mathbf{A}}(z), \hat{\mathbf{B}}(z)$ の収束半径を, それぞれ r_A, r_B とし, 以下の仮定をおく.

仮定 4.1 $r := \min(r_A, r_B) > 1$.

ついで, 次の仮定をおく.

仮定 4.2 以下を満たす非負関数 f が存在する.

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{\overline{\mathbf{A}}(N)\mathbf{e}}{r^{-N}f(N)} &= \mathbf{c}_A, \\ \lim_{N \rightarrow \infty} \frac{\overline{\mathbf{B}}(N)\mathbf{e}}{r^{-N}f(N)} &= \mathbf{c}_B, \\ \lim_{N \rightarrow \infty} \frac{f(N)}{f(N-1)} &= 1.\end{aligned}$$

ただし, $\mathbf{c}_A$ と $\mathbf{c}_B$ はそれぞれ M 次元, M_0 次元の列ベクトルであり, 少なくともどちらか一方は非ゼロベクトルである.

注意 4.3 仮定 4.1 より, $r_A > r_B$ のとき, $\mathbf{c}_A = \mathbf{0}$ であり, $r_A < r_B$ のとき, $\mathbf{c}_B = \mathbf{0}$ である.

これらの仮定のもとで次の定理を導くことができる.

定理 4.4 仮定 2.1, 2.4 および 4.1 が満たされるとき, 以下が成り立つ.

$$\limsup_{N \rightarrow \infty} \frac{\pi^{(N)} - \pi}{r^{-N}f(N)} \leq \left\{ \frac{r}{-\sigma} + 2(r-1)\xi \right\} \{ \bar{\pi}(0)\mathbf{c}_A + \pi(0)\mathbf{c}_B \} \pi, \quad (4.1)$$

$$\liminf_{N \rightarrow \infty} \frac{\pi^{(N)} - \pi}{r^{-N}f(N)} \geq \left\{ \frac{r}{-\sigma} - 2(r-1)\xi \right\} \{ \bar{\pi}(0)\mathbf{c}_A + \pi(0)\mathbf{c}_B \} \pi. \quad (4.2)$$

ただし, $\xi = \|(I - A - \beta_A g)^{-1} \mathbf{e}\|_\infty$ (最大値ノルム) とする.

定理 4.4 から, 次の系を得る.

系 4.5 仮定 2.1, 2.4, 4.1 が満たされているとする. さらに, 任意の $m \in \mathbb{Z}_+$ に対して, $\tilde{\mathbf{A}}(m) = \overline{\mathbf{A}}(m)$ および $\tilde{\mathbf{B}}(m) = \overline{\mathbf{B}}(m)$ であるならば,

$$\lim_{N \rightarrow \infty} \frac{\pi^{(N)} - \pi}{r^{-N}f(N)} = \frac{r\{ \bar{\pi}(0)\mathbf{c}_A + \pi(0)\mathbf{c}_B \}}{-\sigma} \pi, \quad k \in \mathbb{Z}_+, \quad (4.3)$$

が成り立つ.

参考文献

- [1] S. Asmussen. *Applied Probability and Queues*. Springer, New York, Second edition, 2003.
- [2] A. Baiocchi. Analysis of the loss probability of the MAP/G/1/K queue Part I: Asymptotic theory. *Stochastic Models*, 10(4):867–893, 1994.
- [3] Levent Gün. Experimental results on matrix-analytical solution techniques-extensions and comparisons. *Stochastic Models*, 5(4):669–682, 1989.
- [4] Jeongsim Kim and Bara Kim. Asymptotic analysis for loss probability of queues with finite GI/M/1 type structure. *Queueing Systems*, 57(1):47–55, 2007.
- [5] D. M. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Stochastic Models*, 7(1):1–46, 1991.
- [6] H. Masuyama. Error bounds for augmented truncations of discrete-time block-monotone Markov chains under geometric drift conditions. *Advances in Applied Probability*, 47(1):83–105, 2015.
- [7] H. Masuyama. Error bounds for augmented truncations of discrete-time block-monotone Markov chains under subgeometric drift conditions. *SIAM Journal on Matrix Analysis and Applications*, 37(3):877–910, 2016.
- [8] H. Masuyama. A sufficient condition for the subexponential asymptotics of GI/G/1-type Markov chains with queueing applications. *Annals of Operations Research*, 247(1):65–95, 2016.

- [9] H. Masuyama. Continuous-time block-monotone Markov chains and their block-augmented truncations. *Linear Algebra and its Applications*, 514(1):105–150, 2017.
- [10] H. Masuyama. Error bounds for last-column-block-augmented truncations of block-structured Markov chains. *Journal of the Operations Research Society of Japan*, 60(3):271–320, 2017.
- [11] H. Masuyama, T. Kimura, and Y. Katsumata. A subgeometric convergence formula for finite-level M/G/1-type Markov chains via the Poisson equation of the deviation matrix. Preprint arXiv:1809.03179, 2018, 2018.
- [12] M. Miyazawa, Y. Sakuma, and S. Yamaguchi. Asymptotic behaviors of the loss probability for a finite buffer queue with QBD structure. *Stochastic Models*, 23(1):79–95, 2007.
- [13] M. F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. Marcel Dekker, New York, 1989.
- [14] V. Ramaswami. A stable recursion for the steady state vector in Markov chains of M/G/1 type. *Stochastic Models*, 4(1):183–188, 1988.
- [15] H. Schellhaas. On Ramaswami’s algorithm for the computation of the steady state vector in Markov chains of M/G/1-type. *Stochastic Models*, 6:541–550, 1990.

待ち行列モデルを用いた宅配ボックスサービスの性能解析

日出山 慎人¹ Phung-Duc Tuan² 岡田 幸彦²

¹ 筑波大学大学院 システム情報工学研究科 ² 筑波大学 システム情報系

概要：本論文では、宅配ボックスサービスを含む宅配サービスシステムについて、準出生死滅過程を利用して待ち行列モデル化し、解析を行った。そして性能評価指標として宅配ボックスサービスのブロッキング確率、待ち行列内の平均荷物数を定義し、導出した。また、各性能評価指標の数値実験では、待ち行列モデル化による解析結果と乱数を用いたシミュレーションの結果を同時に確認した。そして今回の数値実験結果より、宅配ボックス内荷物の回収に対する受け手の積極性や宅配ボックスサービスの利用率が各性能評価指標に与える影響を明らかにした。

1. はじめに

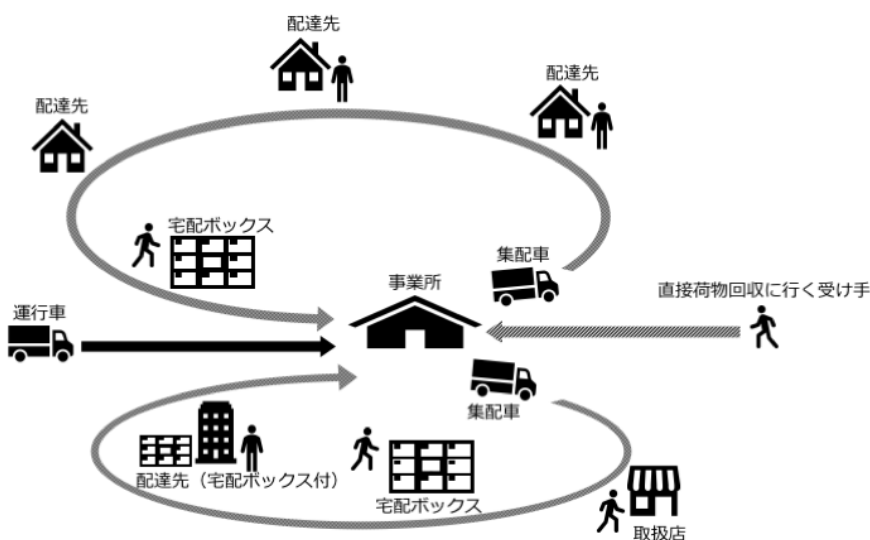
EC 事業の拡大に伴い、国内の宅配便取扱個数は成長を続けている。国土交通省 [1] によると、2017 年度の宅配便取扱個数は約 42.5 億個にものぼり、過去 10 年で 3 割以上も増加している。宅配便取扱個数が増加する一方で、受け手の不在による再配達も発生しており、平成 30 年 4 月期の宅配便再配達率は約 15.0% となった [2]。また、再配達の発生に伴い、再配達に費やすドライバーの労働力や二酸化炭素の発生等の社会的損失も生じている [1]。このような再配達問題への対策案の 1 つとして、宅配ボックスの活用が推進されている。宅配ボックスとは、配送業者が宅配用荷物を格納しておくことで後日受け手が任意のタイミングで荷物を回収できるような設備を指し、駅やマンション等に多く設置されている。宅配ボックスについては様々な研究が行われている。Iwan et al. [3] は宅配ボックスの再配置について、その前後における平均配達荷物数を比較した。Lachapelle et al. [4] は宅配ボックスの立地環境に着目した分析を行った。また、宅配ボックスの実証実験も行われており、福井県あわら市では戸建て用宅配ボックスの設置によって再配達率が 49% から 8% まで減少し [5]、京都府京都市ではアパート用・公共用宅配ボックスの設置によって再配達率が 43% から 15% まで減少した [6]。

しかし、このような研究の中で、宅配ボックスが利用できなくなる状況およびその確率についてはあまり言及されていない。実際の宅配ボックスは大きさが限られているため、同時に格納できる最大荷物数は有限となる。そのため、多くの受け手が同時に宅配ボックスを利用したい場合や受け手の荷物回収速度が遅い場合、宅配ボックスはすぐに満杯となり、それ以上新しい荷物を格納できず、そのタイミングで新しく宅配ボックスサービスを利用したい受け手はサービスを利用できない。ここで、宅配ボックスサービスを利用しようとするタイミングでサービスを利用できない現象をブロッキングと呼び、ブロッキングされる確率をブロッキング確率と呼ぶことにする。ブロッキングは実際にも発生しており、[5] では実証実験中に宅配ボックスが機能しなかった場面の原因として挙げられている。そのため、宅配ボックスサービスのブロッキング確率について示すことには重要な意味があると考えられる。ここで、サービスシステムのブロッキング確率については待ち行列理論を用いたモデル化・解析を行うことで導出可能となる [7]。待ち行列理論を用いた解析は交通・物流分野の研究においても既に応用されている [8, 9]。そのうち、日出山ら [9] では宅配ボックスサービスを含む宅配サービスシステムのモデル化が行われたが、宅配ボックスで荷物受取を希望する受け手の存在を考慮できておらず、さらに性能評価指標については明らかにされていない。

以上を踏まえ、今回の我々の研究目的として、日出山ら [9] を拡張した宅配ボックスサービスを含む宅配サービスシステムについて待ち行列理論を用いた確率モデル化・解析を行う。そして宅配ボックスサービスのブロッキング確率をはじめとする性能評価指標をいくつか定義し、導出する。さらに数値計算・シミュレーションを行い、各性能評価指標の変化および解析結果の妥当性を確認する。

本論文の構成は以下の通りである。第 2 章では、宅配サービスシステムについて整理し、待ち行列モデルで表現する。第 3 章では、第 2 章で待ち行列モデル化したシステムについて、待ち行列理論を用いた解析を行い、安定条件・定常状態確率・性能評価指標を導出する。第 4 章では、第 3 章で定義した安定条件と性能評価指標について数値計算・シミュレーションを行った数値実験例の結果を示す。第 5 章では、結論と今後の課題を述べる。

本章では宅配サービスシステムのモデル化を行う。実際の宅配サービスシステムは図1のように非常に複雑な仕組みをしており、モデル化が難しい。そこで今回は宅配サービスシステムのうち本質的な要素のみを取り入れた宅配サービスシステムについてモデル化を行う。



今回想定する宅配サービスシステムについて、配送業者が管理するような事業所に着目し、その事業所は周辺地域に対して荷物配達を行うものとする。また、配達圏内には大型の宅配ボックスが1つ存在するものとする。事業所には外部から配達依頼の荷物が到着し、事業所は宅配トラックによる配達、宅配ボックスによる配達のどちらかに振り分けてそれぞれサービスを行う。このとき、荷物数をジョブ数として捉え、宅配トラックによる配達用の荷物の待ち行列、宅配ボックスによる配達用の荷物の待ち行列を考えることにより、今回の宅配サービスシステムは2つの待ち行列を含む待ち行列ネットワークとなる。以下では今回の宅配サービスシステムについて、それぞれ荷物の到着・宅配トラック用荷物の待ち行列・宅配ボックス用荷物の待ち行列に分けて整理する。

まずはじめに、事業所には外部から配達依頼の荷物がパラメータ λ のポアソン過程に従って到着する。このとき、実際の EC サイトでは商品購入時に受取方法を指定できるサービスを提供しているものが多い。例として楽天株式会社が提供する楽天市場 [10] では、商品購入時に宅配ボックスへの配送を指定できたり、不在配達時には宅配ボックスに格納するよう予め指定できる。以上を踏まえ、ここでは事業所に到着する荷物について「自宅で受取希望」「宅配ボックスで受取希望」の 2 種類の存在を仮定し、新しく事業所に到着した荷物には確率 $1-r$ で「自宅で受取希望」、確率 r で「宅配ボックスで受取希望」が指定されているものとする。続いて事業所は到着した荷物を宅配トラック用荷物の待ち行列、宅配ボックス用荷物の待ち行列のどちらかに振り分ける。宅配トラック用荷物の待ち行列に振り分けられた荷物は、トラックによる宅配によって各受け手の自宅まで配達される。一方で宅配ボックス用荷物の待ち行列に振り分けられた荷物は、配送業者が宅配ボックスに荷物を格納しておき、その後受け手が荷物を回収する事で配達完了する。ただし、宅配トラック用荷物の待ち行列は無限バッファとし、宅配ボックス用荷物の待ち行列は有限バッファとする。このとき、「自宅で受取希望」の荷物は全て宅配トラック用荷物の待ち行列に振り分けられるが、「宅配ボックスで受取希望」の荷物は宅配ボックス用荷物の待ち行列が満杯でなければ宅配ボックス用荷物の待ち行列へ、満杯ならばサービスを利用できないため仕方なく宅配トラック用荷物の待ち行列へ振り分けられる。

宅配トラック用荷物の待ち行列に振り分けられた荷物についてはトラックを用いて受け手の自宅まで配達を行う。今回のモデルでは宅配トラックは1台のみとし、荷物の積み込み時間は考慮しないものとする。荷物配達は1つの配達先に対して1つずつ行われる。また、ある配達先を出発してから次の

配達先まで到着し、そこで配達作業を終えるまでにかかる時間はパラメータ μ_1 の指数分布に従うものとする。そしてある配達先を訪問した時、受け手が在宅の確率を p 、不在の確率を $1-p$ とする。受け手が在宅の場合、荷物を直接受け手に渡す事で配達成功となり、宅配トラック用荷物数が1減る。受け手が不在の場合、宅配ボックス用荷物の待ち行列が満杯でなければ、荷物は宅配ボックス用荷物の待ち行列に振り分けられ、宅配トラック用荷物数が1減り、宅配ボックス用荷物数が1増える。一方で満杯ならば、荷物はそのまま事業所へ持ち帰られ、各荷物数は変化しない。

2.3. 宅配ボックス用荷物の待ち行列

宅配ボックス用荷物の待ち行列に振り分けられた荷物については宅配ボックスを活用した配達を行う。宅配ボックスには最大 c 個まで荷物を格納でき、事業所には最大 $K-c$ 個まで宅配ボックス用荷物を待機させることができるものとする。このとき、 $c \leq K$ とする。宅配ボックスに格納された荷物は受け手によってそれぞれ独立に1つずつ回収され、1つあたりの回収時間はパラメータ μ_2 の指数分布に従うものとする。また、事業所から宅配ボックスまでの移動時間、宅配ボックスへの荷物格納にかかる時間については今回は考慮しないものとする。

2.4. 待ち行列モデル

時刻 $t \geq 0$ における宅配トラック用の荷物数、宅配ボックス用の荷物数をそれぞれ $i(t), j(t)$ とする。ただし、 $i(t)$ は整数かつ $i(t) \geq 0$, $j(t)$ は整数かつ $0 \leq j(t) \leq K$ を満たすような値とする。このとき、今回想定する宅配サービスシステムは図2のような待ち行列モデルで表現できる。

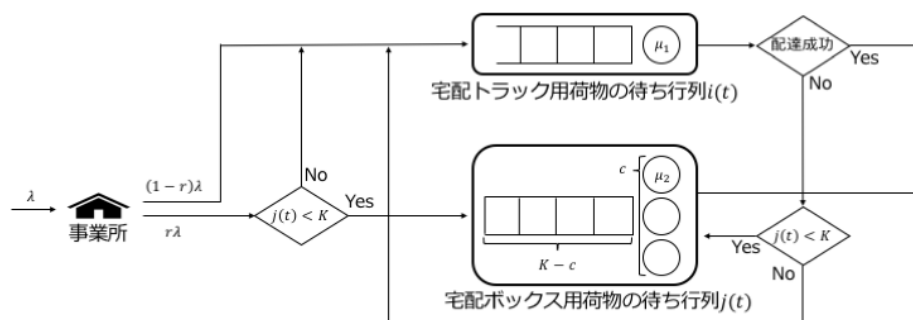


図2: 待ち行列モデル。

3. 解析

本章では、第2章で作成した待ち行列モデルについて、確率過程を用いた解析を進める。

3.1. 準出生死滅過程

第2章で、時刻 $t \geq 0$ における宅配トラック用の荷物数を $i(t)$ 、宅配ボックス用の荷物数を $j(t)$ とした。ここで集合 I, J をそれぞれ $I = \{0, 1, 2, \dots\}, J = \{0, 1, 2, \dots, K\}$ で定義する。このとき、 $S = I \times J$ とすると、確率過程 $\{(i(t), j(t)) \mid t \geq 0\}$ は状態空間 S を持つ連続時間マルコフ連鎖となる。また、一回の推移で $i(t)$ は高々1しか増減しない為、 $\{(i(t), j(t)) \mid t \geq 0\}$ は準出生死滅過程となる。さらに滝根 [11] より、状態空間 S を次のように分解して考える。

$$S = \bigcup_{k=0}^{\infty} \mathcal{L}_k,$$

ただし、 $\mathcal{L}_k := \{(i, j) \mid i = k, j \in J\} (k \in I)$ とする。このとき、今回の準出生死滅過程の無限小生成作用素 Q は次のように表現できる。

$$Q = \begin{matrix} & \mathcal{L}_0 & \mathcal{L}_1 & \mathcal{L}_2 & \mathcal{L}_3 & \mathcal{L}_4 \\ \begin{matrix} \mathcal{L}_0 \\ \mathcal{L}_1 \\ \mathcal{L}_2 \\ \mathcal{L}_3 \\ \mathcal{L}_4 \end{matrix} & \begin{pmatrix} B_0 & A_1 & O & O & O & \cdots \\ A_{-1} & A_0 & A_1 & O & O & \cdots \\ O & A_{-1} & A_0 & A_1 & O & \cdots \\ O & O & A_{-1} & A_0 & A_1 & \cdots \\ O & O & O & A_{-1} & A_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{matrix},$$

ただし,

$$\begin{aligned}
\mathbf{B}_0 =_c & \begin{pmatrix}
0 & 1 & & c-1 & c & & K-1 & K \\
0 & -\lambda & r\lambda & \cdots & 0 & 0 & \cdots & 0 & 0 \\
1 & \mu_2 & -(\mu_2 + \lambda) & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
c-1 & 0 & 0 & \ddots & -\{(c-1)\mu_2 + \lambda\} & r\lambda & \ddots & 0 & 0 \\
& 0 & 0 & \ddots & c\mu_2 & -(c\mu_2 + \lambda) & \ddots & 0 & 0 \\
c+1 & 0 & 0 & \ddots & 0 & c\mu_2 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
K-1 & 0 & 0 & \ddots & 0 & 0 & \ddots & -(c\mu_2 + \lambda) & r\lambda \\
K & 0 & 0 & \cdots & 0 & 0 & \cdots & c\mu_2 & -(c\mu_2 + \lambda)
\end{pmatrix}, \\
\mathbf{A}_{-1} =_c & \begin{pmatrix}
0 & 1 & & c-1 & c & & K-1 & K \\
0 & p\mu_1 & (1-p)\mu_1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
1 & 0 & p\mu_1 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
c-1 & 0 & 0 & \ddots & p\mu_1 & (1-p)\mu_1 & \ddots & 0 & 0 \\
& 0 & 0 & \ddots & 0 & p\mu_1 & \ddots & 0 & 0 \\
c+1 & 0 & 0 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
K-1 & 0 & 0 & \ddots & 0 & 0 & \ddots & p\mu_1 & (1-p)\mu_1 \\
K & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & p\mu_1
\end{pmatrix}, \\
\mathbf{A}_0 =_c & \begin{pmatrix}
0 & 1 & & c-1 & c & & K-1 & K \\
0 & -(\mu_1 + \lambda) & r\lambda & \cdots & 0 & 0 & \cdots & 0 & 0 \\
1 & \mu_2 & -(\mu_1 + \mu_2 + \lambda) & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
c-1 & 0 & 0 & \ddots & -\{\mu_1 + (c-1)\mu_2 + \lambda\} & r\lambda & \ddots & 0 & 0 \\
& 0 & 0 & \ddots & c\mu_2 & -(\mu_1 + c\mu_2 + \lambda) & \ddots & 0 & 0 \\
c+1 & 0 & 0 & \ddots & 0 & c\mu_2 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
K-1 & 0 & 0 & \ddots & 0 & 0 & \ddots & -(\mu_1 + c\mu_2 + \lambda) & r\lambda \\
K & 0 & 0 & \cdots & 0 & 0 & \cdots & c\mu_2 & -(p\mu_1 + c\mu_2 + \lambda)
\end{pmatrix}, \\
\mathbf{A}_1 =_c & \begin{pmatrix}
0 & 1 & & c-1 & c & & K-1 & K \\
0 & (1-r)\lambda & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
1 & 0 & (1-r)\lambda & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
c-1 & 0 & 0 & \ddots & (1-r)\lambda & 0 & \ddots & 0 & 0 \\
& 0 & 0 & \ddots & 0 & (1-r)\lambda & \ddots & 0 & 0 \\
c+1 & 0 & 0 & \ddots & 0 & 0 & \ddots & 0 & 0 \\
& \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
K-1 & 0 & 0 & \ddots & 0 & 0 & \ddots & (1-r)\lambda & 0 \\
K & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \lambda
\end{pmatrix}.
\end{aligned}$$

このとき, $\mathbf{B}_0, \mathbf{A}_{-1}, \mathbf{A}_0, \mathbf{A}_1, \mathbf{O}$ はいずれも $(K+1)$ 次正方行列である (ただし, $\mathbf{O}$ は零行列とする).

3.2. 安定条件

続いて、安定条件を導出する。今回の安定条件の導出は滝根 [11] の方法を参考に進める。はじめに、以下のような無限小生成作用素 $\mathbf{A}$ を定義する。

$$\mathbf{A} = \mathbf{A}_{-1} + \mathbf{A}_0 + \mathbf{A}_1,$$

$$\begin{matrix} & 0 & & c-1 & & c & & K \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ c-1 \\ c \\ \vdots \\ K-1 \\ K \end{matrix} & \begin{pmatrix} -\{(1-p)\mu_1 + r\lambda\} & \cdots & 0 & & 0 & \cdots & 0 \\ \mu_2 & \ddots & 0 & & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \ddots & -\{(1-p)\mu_1 + (c-1)\mu_2 + r\lambda\} & & (1-p)\mu_1 + r\lambda & \ddots & 0 \\ 0 & \ddots & c\mu_2 & & -((1-p)\mu_1 + c\mu_2 + r\lambda) & \ddots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & & 0 & \ddots & (1-p)\mu_1 + r\lambda \\ 0 & \cdots & 0 & & 0 & \cdots & -c\mu_2 \end{pmatrix} \end{matrix}.$$

このような $\mathbf{A}$ に対し、以下の等式を満たす、同質な部分における相の定常分布 $\boldsymbol{\eta}$ を考える。

$$\boldsymbol{\eta}\mathbf{A} = \mathbf{0}, \quad \boldsymbol{\eta}\mathbf{e} = 1,$$

ただし、 $\mathbf{0}$ は全要素が 0 の適切な次元数の行ベクトル、 $\mathbf{e}$ は全要素が 1 の適切な次元数の列ベクトルとする。ここで、 $\mathbf{A}$ は到着率 $(1-p)\mu_1 + r\lambda$ 、サービス率 μ_2 の M/M/c/K 待ち行列に対する無限小生成作用素となっている事に注目する。したがって定常分布 $\boldsymbol{\eta}$ は式 (1) で表現される $(K+1)$ 次元の行ベクトルとなる。

$$\boldsymbol{\eta} = \left(\frac{0}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k}, \quad \cdots, \quad \frac{\frac{\rho^c}{c!}}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k}, \quad \cdots, \quad \frac{\frac{\rho^c}{c!} \left(\frac{\rho}{c}\right)^{K-c}}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k} \right), \quad (1)$$

ただし、 $\rho = \frac{(1-p)\mu_1 + r\lambda}{\mu_2}$ とする。また、同質な部分における相の定常分布 $\boldsymbol{\eta}$ が式 (1) で与えられる時、滝根 [11] より、安定条件は式 (2) で与えられる。

$$\boldsymbol{\eta}\mathbf{A}_1\mathbf{e} - \boldsymbol{\eta}\mathbf{A}_{-1}\mathbf{e} < 0. \quad (2)$$

このとき、

$$\boldsymbol{\eta}\mathbf{A}_1\mathbf{e} - \boldsymbol{\eta}\mathbf{A}_{-1}\mathbf{e} = \boldsymbol{\eta}(\mathbf{A}_1 - \mathbf{A}_{-1})\mathbf{e},$$

$$\begin{aligned} &= \boldsymbol{\eta} \begin{pmatrix} (1-r)\lambda - p\mu_1 & -(1-p)\mu_1 & \cdots & 0 & 0 \\ 0 & (1-r)\lambda - p\mu_1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & (1-r)\lambda - p\mu_1 & -(1-p)\mu_1 \\ 0 & 0 & \cdots & 0 & \lambda - p\mu_1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \\ &= \boldsymbol{\eta} \begin{pmatrix} (1-r)\lambda - \mu_1 \\ (1-r)\lambda - \mu_1 \\ \vdots \\ (1-r)\lambda - \mu_1 \\ \lambda - p\mu_1 \end{pmatrix}, \\ &= (1-r)\lambda - \mu_1 + \{r\lambda + (1-p)\mu_1\} \frac{\frac{\rho^c}{c!} \left(\frac{\rho}{c}\right)^{K-c}}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k}, \\ &< 0. \end{aligned}$$

以上より、今回のモデルにおける安定条件として式 (3) を得る.

$$(1-r)\lambda + \{r\lambda + (1-p)\mu_1\} \frac{\frac{\rho^c}{c!} \left(\frac{\rho}{c}\right)^{K-c}}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k} < \mu_1. \quad (3)$$

式 (3) の左辺について、第一項は「自宅で受取希望」の荷物の到着率を表現し、第二項は同質な部分において宅配ボックスサービスにブロッキングされた荷物の到着率を表現している. したがって、式 (3) の安定条件が満たされるためには、事業所のサービス率がこれらの到着率の和を上回る必要がある.

3.3. 定常状態確率

続いて、安定条件の下で定常状態確率を計算する. 状態 (i, j) の定常状態確率を $\pi_{i,j}$ とし、以下のように定義する.

$$\pi_{i,j} = \lim_{t \rightarrow \infty} P(i(t) = i, j(t) = j).$$

そして π を以下のように定義する.

$$\pi = \left(\pi_0, \pi_1, \pi_2, \dots \right),$$

ただし,

$$\pi_i = \left(\pi_{i,0}, \pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,K} \right).$$

この時、大域平衡方程式 $\pi Q = 0$ より式 (4) および式 (5) を得る.

$$\pi_0 B_0 + \pi_1 A_{-1} = 0, \quad (4)$$

$$\pi_{k-1} A_1 + \pi_k A_0 + \pi_{k+1} A_{-1} = 0, \quad k = 1, 2, 3, \dots \quad (5)$$

ここから滝根 [11] により、定常状態確率および率行列 R について、式 (6) および式 (7) を得る.

$$\pi_k = \pi_1 R^{k-1}, \quad k = 1, 2, \dots, \quad (6)$$

$$A_1 + R A_0 + R^2 A_{-1} = 0. \quad (7)$$

率行列 R は、式 (7) を数值的に解く事で得られる. 詳細な解法については滝根 [11] を参照されたい. また、この方法で得られた率行列 R について、式 (3) の安定条件を満たしているならば、最大固有値は 1 未満となり、式 (8) を満たす.

$$\sum_{k=1}^{\infty} \pi_k e = \pi_1 (I - R)^{-1} e < \infty, \quad (8)$$

ただし、 I は適切な次元数の単位行列とする.

続いて π_0, π_1 を導出する. まずは式 (4) および $k = 1$ とした式 (5) を併せて行列表現することで、式 (9) を得る.

$$\begin{pmatrix} \pi_0 & \pi_1 \end{pmatrix} \begin{pmatrix} B_0 & A_1 \\ A_{-1} & A_0 + R A_{-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}. \quad (9)$$

また、定常状態確率の総和が 1 になる性質より式 (10) を得る.

$$\sum_{k=0}^{\infty} \pi_k e = 1. \quad (10)$$

そして式 (6)、式 (8) および式 (10) より、式 (11) を得る.

$$\pi_0 e + \pi_1 (I - R)^{-1} e = 1. \quad (11)$$

以上より、滝根 [11] の方法より率行列 R が計算でき、式 (9) および式 (11) を解く事で π_0 および π_1 が計算でき、式 (6) を解く事で π_k ($k = 2, 3, \dots$) が計算できる.

3.4. 性能評価指標

続いて、性能評価指標を定義し、導出する．今回は性能評価指標として宅配ボックスサービスのブロッキング確率、宅配トラック用荷物の平均数、宅配ボックス用荷物の平均数を定義する．ここでブロッキング確率とは、事業所に到着した「宅配ボックスで受取希望」の荷物または宅配に失敗した荷物が宅配ボックスサービスを利用したいタイミングで宅配ボックス用荷物の待ち行列が満杯であり、サービスを受けられない確率とする．

まずはじめに、ブロッキング確率 P_b を導出する．今回の宅配ボックスサービスのブロッキング確率は式 (12) のように表現できる．

$$P_b = \frac{r\lambda\pi_0\acute{e} + \{(1-p)\mu_1 + r\lambda\} \sum_{i=1}^{\infty} \pi_i\acute{e}}{r\lambda\pi_0e + \{(1-p)\mu_1 + r\lambda\} \sum_{i=1}^{\infty} \pi_ie}, \quad (12)$$

ただし、 $\acute{e}$ は最後の要素の値が 1 でそれ以外の要素の値が 0 の適切な次元数の列ベクトルとする．そして式 (8) および式 (12) より、 P_b について式 (13) を得る．

$$P_b = \frac{r\lambda\pi_0\acute{e} + \{(1-p)\mu_1 + r\lambda\}\pi_1(I - R)^{-1}\acute{e}}{r\lambda\pi_0e + \{(1-p)\mu_1 + r\lambda\}\pi_1(I - R)^{-1}e}. \quad (13)$$

次に宅配トラック用荷物の平均数 N_1 を導出する．まず期待値の定義より、 N_1 は式 (14) のように表現される．

$$N_1 = E[i(t)] = \sum_{i=0}^{\infty} i\pi_ie. \quad (14)$$

また、期待値に関する数学的な性質により、式 (15) を得る．

$$E[i(t)] = \sum_{n=1}^{\infty} P(i(t) \geq n) = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \pi_ke. \quad (15)$$

式 (6) を式 (15) に代入することで、式 (16) を得る．

$$E[i(t)] = \pi_1 \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} R^{k-1}e. \quad (16)$$

また、式 (8) を繰り返し利用することにより、式 (17) を得る．

$$\sum_{n=1}^{\infty} \sum_{k=n}^{\infty} R^{k-1} = \{(I - R)^{-1}\}^2. \quad (17)$$

ここで式 (14)、式 (15) および式 (17) より、 N_1 について式 (18) を得る．

$$N_1 = \pi_1 \{(I - R)^{-1}\}^2 e. \quad (18)$$

最後に宅配ボックス用荷物の平均数 N_2 を導出する．まず期待値の定義より、 N_2 は式 (19) のように表現される．

$$N_2 = E[j(t)] = \sum_{i=0}^{\infty} \pi_i\tilde{e} = \pi_0\tilde{e} + \sum_{i=1}^{\infty} \pi_i\tilde{e}, \quad (19)$$

ただし、 $\tilde{e}$ は n 番目の要素の値が $n - 1$ の適切な次元数の列ベクトルとする．そして式 (8) および式 (19) より、 N_2 について式 (20) を得る．

$$N_2 = \pi_0\tilde{e} + \pi_1(I - R)^{-1}\tilde{e}. \quad (20)$$

4. 数値実験と考察

本章では、第 3.2 節で導出した安定条件および第 3.4 節で導出した各性能評価指標に関する数値実験例を示し、考察を行う．

4.1. 安定領域

まずはじめに、第3.2節で導出した安定条件に関する数値実験を行う。ここでは宅配ボックスの設置がもたらす安定領域の増分を示す。ただし、安定領域とは安定条件が満たされている領域とする。今回は μ_1, λ をそれぞれ横軸、縦軸とするグラフにおいて安定領域を確認するため、安定領域は $\mu_1, \lambda \geq 0$ および式(3)を満たす領域とする。このとき、式(21)は安定領域の境界となる。

$$(1-r)\lambda + \{r\lambda + (1-p)\mu_1\} \frac{\frac{\rho^c}{c!} \left(\frac{\rho}{c}\right)^{K-c}}{\sum_{k=0}^c \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \sum_{k=1}^{K-c} \left(\frac{\rho}{c}\right)^k} = \mu_1. \quad (21)$$

ここで、ある状況において宅配ボックスを設置した場合の安定領域について考えたい。設置された宅配ボックスが十分に機能した場合、宅配に失敗した荷物は宅配ボックス用荷物に回すことで再配達を回避できるため、宅配が確率1で成功する状況に近い安定領域が得られる。一方で、設置された宅配ボックスが十分に機能しない場合、事業所は再配達を行わなければならない、宅配ボックスが設置されていない状況に近い安定領域が得られる。以上を踏まえ、各パラメータが $r, \lambda, \mu_1, \mu_2, p$ で与えられた状況において固定された c, K の宅配ボックスサービスの導入が安定領域に与える影響を評価する方法として、次の3つの場合それぞれの安定領域の境界を考える。

(a) 宅配成功率が1で、宅配ボックスを設置する場合（すなわち、 $p = 1$ に変更した場合）

(b) 宅配成功率が p で、宅配ボックスを設置する場合（与えられた状況）

(c) 宅配成功率が p で、宅配ボックスを設置しない場合（すなわち、 $c = K = 0$ に変更した場合）

ここで、宅配ボックスが十分に機能した場合、(b)の安定領域の境界は(a)の安定領域の境界に漸近し、宅配ボックスが十分に機能しなかった場合、(b)の安定領域の境界は(c)の安定領域の境界に漸近する。そして(b)の安定領域の境界と(c)の安定領域の境界に挟まれる領域が、宅配ボックスサービスの導入がもたらす安定領域の増分となる。

安定領域の数値実験として、 $r = 0.10, p = 0.85$ で与えられた状況において $c = 50, K = 100$ の宅配ボックスサービスの導入がもたらす安定領域の増分を考え、 $\mu_2 = 0.1, 1.0, 5.0$ それぞれの場合について検証した。ここで μ_2 は受け手による宅配ボックス内荷物の回収時間のパラメータであり、 μ_2 が大きいほど受け手が荷物回収に対して積極的であると考えられる。そして結果は図3のようになった。

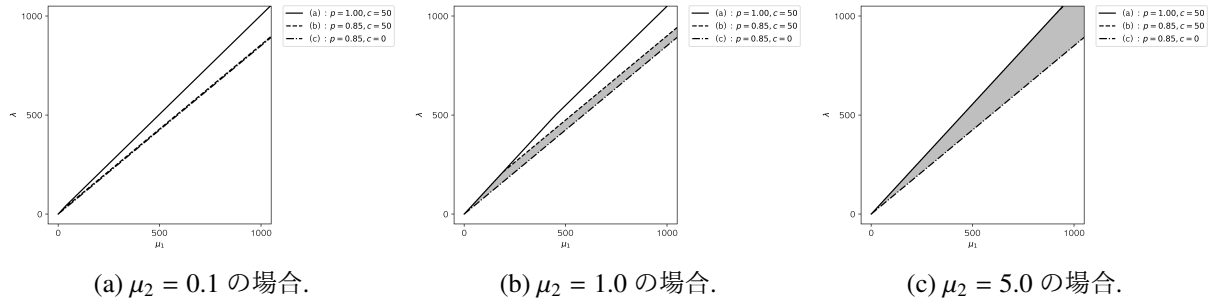


図3: 安定領域の増分の数値実験結果.

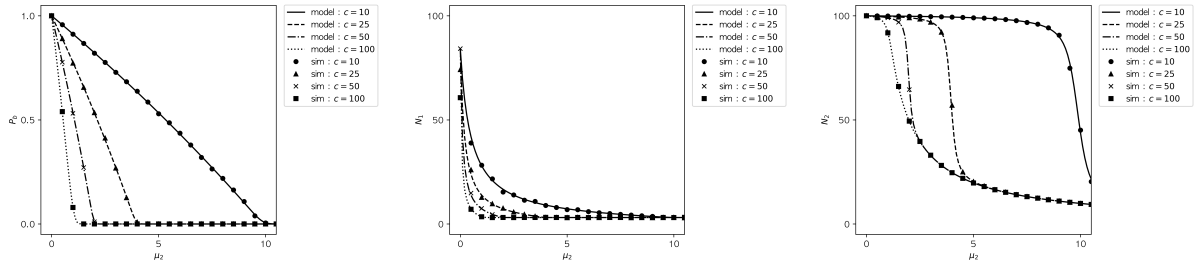
ここで、今回の宅配ボックスサービスの導入がもたらす安定領域の増分について色付けを行った。この結果より、 μ_2 が大きいほど安定領域の増分は大きいことがわかる。また、図3aでは(b)が(c)にかなり漸近し、図3cでは(b)が(a)にかなり漸近していることがわかる。すなわち、受け手が宅配ボックス内荷物の回収に積極的ならば安定条件は大きく広がり、直感的にも納得しやすい結果となった。

4.2. 性能評価指標

続いて各性能評価指標に関する数値実験を行う。各性能評価指標の数値実験は μ_2, r それぞれを変化させた場合について行った。

まず1つ目の数値実験例として、 $r = 0.10, \lambda = 420, \mu_1 = 500, K = 100, p = 0.85$ とし、 $c = 10, 25, 50, 100$ それぞれの場合について、 μ_2 を変化させて検証した。そして結果は図4のようになった。

図4より、各性能評価指標は μ_2 の増加に対して減少傾向が示され、特に c が大きい場合に各性能評価指標の減少量は大きくなった。また、図4bにて、 $\mu_2 = 0$ は宅配ボックスが設置されていない状況

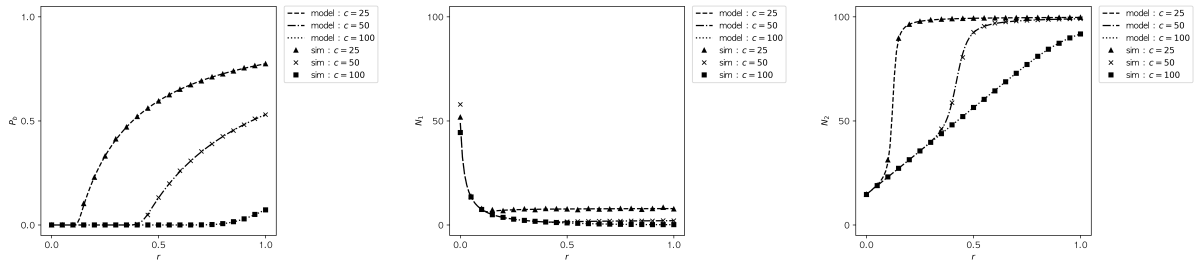


(a) ブロッキング確率. (b) 宅配トラック用荷物の平均数. (c) 宅配ボックス用荷物の平均数.

図 4: μ_2 を操作した場合の各性能評価指標の数値実験結果.

とほぼ同値であるが, μ_2 の増加によって N_1 は大幅に減少していることがわかる. そして図 4c では, 各 c に対し, ある μ_2 付近から N_2 が急落していることがわかる. 以上より, μ_2 の増加に対しては, P_b, N_1, N_2 はいずれも減少することがわかった.

続いて, $\lambda = 490, \mu_1 = 500, \mu_2 = 5.0, K = 100, p = 0.85$ とし, $c = 25, 50, 100$ それぞれの場合について, r を変化させて検証した. そして結果は図 5 のようになった.



(a) ブロッキング確率. (b) 宅配トラック用荷物の平均数. (c) 宅配ボックス用荷物の平均数.

図 5: r を操作した場合の各性能評価指標の数値実験結果.

まず図 5a より, r の増加に対してブロッキング確率も増加した. これについて, r の増加は宅配ボックスサービスの利用率の増加を意味しており, 宅配ボックスサービスの利用が活発になったことで宅配ボックスサービスのブロッキング確率も上昇していることがわかる. 続いて図 5b より, r の増加に対して N_1 はある値まで減少してから再び増加している. これについて, r が低い場合には事業所に負担が集中し, r が高い場合には宅配ボックスに負担が集中するため, 適当な r を選択することで負担がうまく分散されるものと考察できる. 最後に図 5c について, 図 4c の場合と似た振る舞いをしており, 各 c に対し, ある r 付近から N_2 が急伸している. 以上より今回の場合, r の増加に対しては, P_b, N_2 は増加し, N_1 はある値まで減少してから再び増加することがわかった.

5. 結論と今後の課題

本論文では宅配ボックスサービスを含む宅配サービスシステムについて本質的な要素のみを取り入れ, 待ち行列モデルとして表現した. そして準出生死滅過程として待ち行列理論を用いた解析を行い, 安定条件, 定常状態確率, 性能評価指標の導出を行った. 今回は性能評価指標として宅配ボックスサービスのブロッキング確率, 宅配トラック用荷物の平均数, 宅配ボックス用荷物の平均数を定義・導出した. 数値実験では, まず宅配ボックスサービスの導入による安定領域の増分について数値計算を行い, 続いて各性能評価指標の変化について数値計算・シミュレーションを行った. そして数値実験結果から, 宅配ボックス内荷物の回収に対して受け手が積極的なほど安定領域は大きく広がり, 各性能評価指標も改善された. しかし, 宅配ボックスサービスの利用率の増加に対しては今回, 宅配ボックスのブロッキング確率および宅配ボックス用荷物の平均数は増加傾向であるのに対し, 宅配トラック用荷物の平均数はある値まで減少してから再び増加することがわかった.

先行研究では, 宅配ボックスの再配置による配達数の変化に関する研究や [3], 宅配ボックスの立地

環境を数値化して分類・分析する研究が行われていた [4]。しかし、これらの研究の中では宅配ボックス利用状況の動的な変化については言及されてなかった。また、実証実験では宅配ボックスの活用によって再配達率は下がることが示されていた [5, 6]。しかし実証実験中には宅配ボックスが機能しなかった場面もあり、その原因としてブロッキングが挙げられていたが、理論的なアプローチは行われていなかった。そして日出山ら [9] は宅配サービスシステムの確率モデル化を行ったが、ブロッキング確率をはじめとする各性能評価指標の定義・導出については今後の課題としていた。そこで、本研究は日出山ら [9] のモデルを拡張した宅配サービスシステムについて確率過程を用いてモデル化し、宅配サービスシステム内荷物数の動的変化の表現を可能にし、さらにブロッキング確率をはじめとする各性能評価指標を理論的に導出した。以上が本研究の理論的な貢献と考えられる。

しかし、実際の宅配サービスシステムは今回のモデルよりもはるかに複雑なシステムとなっている。そのため、より実際の宅配サービスシステムに近いモデルを作る為には更に複雑な要素を追加し、解析しなければならない。また、今回は性能評価指標として宅配ボックスサービスのブロッキング確率および各待ち行列の平均荷物数を定義し、数値実験結果により受け手の荷物回収速度の重要性を示した。しかし今回は宅配ボックスの利用による受け手の負担等を考慮できていないため、このとき実際には受け手の負担が増大化してしまう。そのため、今後は宅配サービスシステムを通して配り手および受け手が負担するコストを性能評価指標として定義し、トレードオフ問題として扱う必要がある。

謝辞

この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構（NEDO）の委託業務の結果得られたものです。

参考文献

- [1] 国土交通省, 宅配便の再配達削減に向けて, http://www.mlit.go.jp/seisakutokatsu/freight/re_delivery_reduce.html, 2018 年 12 月 14 日最終アクセス.
- [2] 国土交通省, 平成 30 年 4 月期の宅配便再配達率について, <http://www.mlit.go.jp/common/001239385.pdf>, 2018 年 12 月 14 日最終アクセス.
- [3] Iwan, S., Kijewska, K., and Lemke, J., Analysis of parcel lockers' efficiency as the last mile delivery solution - the results of the research in Poland, *Transportation Research Procedia*, **12** (2016), 644–655.
- [4] Lachapelle, U., Burke, M., Brotherton, A., and Leung, A., Parcel locker systems in a car dominant city: Location, characterisation and potential impacts on city planning and consumer travel access, *Journal of Transport Geography*, **71** (2018), 1–14.
- [5] パナソニック株式会社, 「宅配ボックス実証実験」最終結果報告, <https://news.panasonic.com/jp/press/data/2017/06/jn170608-1/jn170608-1.html>, 2018 年 12 月 14 日最終アクセス.
- [6] パナソニック株式会社, 「京(みやこ)の再配達を減らそうプロジェクト」再配達率が 43 % から 15 % に減少, <https://news.panasonic.com/jp/press/data/2018/03/jn180319-1/jn180319-1.html>, 2018 年 12 月 14 日最終アクセス.
- [7] Phung-Duc, T., Multiserver queues with finite capacity and setup time, *Analytical and Stochastic Modelling Techniques and Applications*, LNCS 9081 (2015), 173–187.
- [8] Wang, Y., Guo, J., Ceder, A. A., Currie, G., Dong, W., and Yuan, H., Waiting for public transport services: Queueing analysis with balking and reneging behaviors of impatient passengers, *Transportation Research Part B: Methodological*, **63** (2014), 53–76.
- [9] 日出山慎人, Phung-Duc, T., 岡田幸彦, 待ち行列理論を用いた宅配ボックスサービスのモデル化, 日本経営工学会 2018 年秋季大会, 2018 年 10 月 27 日–2018 年 10 月 28 日.
- [10] 楽天株式会社, 【楽天市場】Shopping is Entertainment! : インターネット最大級の通信販売、通販オンラインショッピングコミュニティ, <https://www.rakuten.co.jp/>, 2018 年 12 月 14 日最終アクセス.
- [11] 滝根哲哉, M/M/1 を越えて 一準出生死滅過程への招待一, オペレーションズ・リサーチ, **59** No.4 (2014), 179–184.

マルコフ連鎖における条件付き定常分布の線形不等式系による特徴づけ

木村 雅俊 滝根 哲哉
大阪大学 工学研究科 電気電子情報工学専攻

概要： 本稿では、非負の整数からなる状態空間上で定義されたマルコフ連鎖における、状態が N 以下であるという条件付きの定常分布を考察する。既存研究において、条件付き定常分布を相対的内部に含む凸多面体の性質が論じられている。一般に、凸多面体は線形不等式系の解空間と等価であるため、これらの結果は条件付き定常分布が満たす線形不等式系を通じても考察可能である。本稿では、そのような線形不等式系が大域平衡方程式から直ちに得られることを示した上、既存研究で議論された凸多面体の性質を線形不等式系を通じて再考する。さらに、従来の行列解析法では活用されていないマルコフ連鎖の新たな構造の概念を導入し、これを用いた条件付き定常分布の特徴付けを行う。

1. はじめに

本稿では、可算無限状態空間 $\mathbb{Z}^+ := \{0, 1, \dots\}$ 上で定義される連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考察する。 $\{X(t); t \geq 0\}$ の推移率行列を Q とする。

$$Q = \begin{pmatrix} -q_0 & q_{0,1} & q_{0,2} & \cdots \\ q_{1,0} & -q_1 & q_{1,2} & \cdots \\ q_{2,0} & q_{2,1} & -q_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

推移率行列 Q の (i, j) 要素 $q_{i,j}$ ($i, j \in \mathbb{Z}^+, i \neq j$) は状態 i から状態 j への推移率で与えられ、対角要素 q_i ($i \in \mathbb{Z}^+$) は行和が 0 となるように定められる。以下では $q_i < \infty$ ($i \in \mathbb{Z}^+$) であり、マルコフ連鎖 $\{X(t); t \geq 0\}$ は定常かつエルゴード的であると仮定する。マルコフ連鎖 $\{X(t); t \geq 0\}$ の定常分布を $\pi = (\pi_0 \ \pi_1 \ \cdots)$ とする。ただし、 $\pi_i = \Pr(X(0) = i)$ ($i \in \mathbb{Z}^+$) である。一般に、エルゴード的な連続時間マルコフ連鎖の定常分布 π は大域平衡方程式の唯一解として与えられる。

$$\pi Q = 0, \quad \pi e = 1 \quad (1)$$

ただし、 e は全ての要素が 1 である適当な次元の列ベクトルである。なお、 $\mathbb{Z}^+$ 上で定義された既約で正再帰的な離散時間マルコフ連鎖の定常分布 π_D に興味がある場合には、遷移確率行列を P_D としたとき、 $\pi_D = \pi_D P_D$ を満たすため、推移率行列 $Q = P_D - I$ で特徴づけられる連続時間マルコフ連鎖の定常分布を考察すれば良い。

状態空間が可算無限な場合、大域平衡方程式には可算無限個の未知数 π_i ($i \in \mathbb{Z}^+$) が含まれるため、定常分布 π を厳密に計算することは一般には困難である。これに対して二つのアプローチが良く知られている。1 つ目は切断技法と呼ばれるアプローチである。切断技法では、十分に大きな $N \in \mathbb{Z}^+$ に対して、状態空間 $\mathbb{Z}^+$ を有限の部分集合 $\mathbb{Z}_0^N$ とその補集合 $\mathbb{Z}_{N+1}^\infty$ に分割する。ただし、 $\mathbb{Z}_m^\ell$ ($m, \ell \in \mathbb{Z}^+$) ならびに $\mathbb{Z}_m^\infty$ ($m \in \mathbb{Z}^+$) は以下で与えられる状態集合を表す。

$$\mathbb{Z}_m^\ell = \begin{cases} \{m, m+1, \dots, \ell\}, & m \leq \ell, \\ \emptyset, & m > \ell, \end{cases} \quad \mathbb{Z}_m^\infty = \{m, m+1, \dots\}$$

また、状態空間の分割に合わせて定常分布 π と推移率行列 Q を以下のように分割する。

$$\pi = \left(\overset{\mathbb{Z}_0^N}{\pi^{(1)}(N)} \ \overset{\mathbb{Z}_{N+1}^\infty}{\pi^{(2)}(N)} \right), \quad Q = \overset{\mathbb{Z}_0^N}{\mathbb{Z}_{N+1}^\infty} \begin{pmatrix} \overset{\mathbb{Z}_0^N}{Q^{(1,1)}(N)} & \overset{\mathbb{Z}_{N+1}^\infty}{Q^{(1,2)}(N)} \\ Q^{(2,1)}(N) & Q^{(2,2)}(N) \end{pmatrix} \quad (2)$$

続いて、 $\{X(t); t \geq 0\}$ の近似として $\mathbb{Z}_0^N$ 上で定義された有限状態マルコフ連鎖を構築する。すなわち、 $(N+1)$ 次元の推移率行列 $[Q^{(1,1)}(N) + Q_A(N)]$ を用意する。ただし、 $Q_A(N)$ は $[Q^{(1,1)}(N) + Q_A(N)]e = 0$ を満たす非負行列である。ここで $\pi(N) = (\pi_0(N) \ \pi_1(N) \ \cdots \ \pi_N(N))$ を状態が N 以下という条件付きの定常分布とする。定義より、条件付き定常分布は次式を満たす。

$$\pi(N) = \frac{\pi^{(1)}(N)}{\pi^{(1)}(N)e} \quad (3)$$

切断技法では、推移率行列 $[Q^{(1,1)}(N) + Q_A(N)]$ をもつ有限状態マルコフ連鎖の定常分布を条件付き定常分布 $\pi(N)$ の近似 $\pi^{\text{approx}}(N)$ として採用する。

$$\pi^{\text{approx}}(N)[Q^{(1,1)}(N) + Q_A^{(1,1)}(N)] = \mathbf{0}, \quad \pi^{\text{approx}}(N)e = 1$$

なお、 $Q_A^{(1,1)}(N)$ を適切に選択すれば、 $N \rightarrow \infty$ とした際 $(\pi^{\text{approx}}(N) \mathbf{0})$ は定常分布 π に収束することが知られている。すなわち、 N が十分に大きいとき、 $(\pi^{\text{approx}}(N) \mathbf{0})$ は π の良い近似となることが期待される。

2つ目は、応用において頻出する推移率行列 Q の構造を利用する、行列解析法と呼ばれるアプローチである。行列解析法では、状態空間 $\mathbb{Z}^+$ をレベルと呼ばれる無限個の排反な有限部分集合 $\mathcal{L}_n$ ($n = 0, 1, \dots$) に分割し、任意のレベルからの遷移を考えたとき、上昇方向に飛び越しが無い場合（レベルが上昇するときは常に1レベルずつ: G/M/1型）あるいは下降方向に飛び越しが無い場合（レベルが下降するときは常に1レベルずつ: M/G/1型）について、定常分布の数値計算法が研究されている [2, 3, 4]。切断技法、行列解析法のいずれにおいても、（条件付き）定常分布を線形連立方程式で特徴づけ、それに基づいて数値計算を行うことになる。

これらとは別に、線形不等式系によって（条件付き）定常分布を特徴づけ、それに基づいた数値計算法が最近、提案されている [1, 5, 7]。文献 [1, 5] ではレベル依存する M/G/1 型マルコフ連鎖を対象として、M/G/1 型という構造を最大限活用した数値計算法が与えられている。一方、文献 [7] では、レベル間での飛び越しなしという構造を持たないマルコフ連鎖に対して、条件付き定常分布 $\pi(N)$ がある凸多面体の相対的内部に含まれること、ならびにそれに基づく数値計算法が提案されている。

本稿の目的は、文献 [7] で得られている結果を線形不等式系という観点から再考し、推移率行列の北西角に位置する $(K+1) \times (K+1)$ 部分行列 ($K \geq N$) がもつ条件付き定常分布 $\pi(N)$ に関する全情報を洗い出すことにある。一般に、凸多面体は線形不等式系の解空間とみなすことができる [6, Theorem 2.15]。本稿の前半では、文献 [7] で示された凸多面体が大域平衡方程式から直接的かつ極めて簡単に導出できることを示す。また、推移率行列の北西角に位置する $(N+1) \times (N+1)$ 部分行列と $(K+1) \times (K+1)$ 部分行列 ($K \geq N$) がもつ条件付き定常分布 $\pi(N)$ に関する情報の関係についても線形不等式系という観点から議論する。本稿の後半では、従来の行列解析法では活用できなかった新しい構造を導入し、推移率行列の北西角に位置する $(K+1) \times (K+1)$ 部分行列 ($K \geq N$) がもつ条件付き定常分布に関する最も詳細な情報を同定する。なお、ここで導入する新しい構造は、従来のレベル間の下降方向への飛越なしという概念を自然な形で拡張したものとなっている。

本稿の構成は以下のとおりである。2節において、文献 [7] の結果をまとめる。3節において、文献 [7] の結果を線形不等式系という観点から再考する。4節では新たな構造として skip-free 集合を導入し、推移率行列の北西角が持つ最も詳細な情報を明らかにする。最後に5節で結論を述べる。

2. 既存研究 [7]

本節では、文献 [7] で得られている結果をまとめる。まず、任意の N, K ($N \leq K$) に対して、状態空間 $\mathbb{Z}^+$ を $\mathbb{Z}_0^N, \mathbb{Z}_{N+1}^K$, ならびに $\mathbb{Z}_{K+1}^\infty$ の三つに分割する。

$$\pi = \left(\begin{array}{ccc} \mathbb{Z}_0^N & \mathbb{Z}_{N+1}^K & \mathbb{Z}_{K+1}^\infty \\ \pi^{(1)}(K, N) & \pi^{(2)}(K, N) & \pi^{(3)}(K, N) \end{array} \right) \quad (4)$$

$$Q = \begin{array}{c} \mathbb{Z}_0^N \\ \mathbb{Z}_{N+1}^K \\ \mathbb{Z}_{K+1}^\infty \end{array} \left(\begin{array}{ccc} Q^{(1,1)}(K, N) & Q^{(1,2)}(K, N) & Q^{(1,3)}(K, N) \\ Q^{(2,1)}(K, N) & Q^{(2,2)}(K, N) & Q^{(2,3)}(K, N) \\ Q^{(3,1)}(K, N) & Q^{(3,2)}(K, N) & Q^{(3,3)}(K, N) \end{array} \right) \quad (5)$$

$K = N$ の場合、式 (4)、式 (5) は式 (2) と等価である。さらに、 $d_{i,j}$ ($i, j \in \mathbb{Z}^+, i \neq j$) を次式で定義する。

$$d_{i,j} = \begin{cases} 1, & q_{i,j} > 0 \\ 0, & \text{その他} \end{cases}$$

$d_{i,j}$ は、状態 i から状態 j への直接遷移が可能か否かを表す指示関数である。また、状態空間 $\mathbb{Z}^+$ の任意の空でない真部分集合 $\mathcal{B}$ に対して、 $d(i, \mathcal{B})$ ($i \in \mathbb{Z}^+, i \notin \mathcal{B}$) ならびに $d(\mathcal{B}, j)$ ($j \in \mathbb{Z}^+, j \notin \mathcal{B}$) を次式で定義する。

$$d(i, \mathcal{B}) = \begin{cases} 1, & \sum_{j \in \mathcal{B}} d_{i,j} > 0, \\ 0, & \text{その他,} \end{cases} \quad d(\mathcal{B}, j) = \begin{cases} 1, & \sum_{i \in \mathcal{B}} d_{i,j} > 0 \\ 0, & \text{その他} \end{cases} \quad (6)$$

さらに、これらを用いて $\mathcal{E}(K, N)$ を次式で定義する。

$$\mathcal{E}(K, N) = \{j \in \mathbb{Z}_0^N; [\mathbf{d}^{(1)}(K, N) + \mathbf{d}^{(2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1}\mathbf{Q}^{(2,1)}(K, N)]_j > 0\} \quad (7)$$

ただし、 $\mathbf{d}^{(i)}(K, N)$ ($i = 1, 2$) は以下で与えられる。

$$\begin{aligned} \mathbf{d}^{(1)}(K, N) &= (d(\mathbb{Z}_{K+1}^\infty, 0) \ d(\mathbb{Z}_{K+1}^\infty, 1) \ \cdots \ d(\mathbb{Z}_{K+1}^\infty, N)) \\ \mathbf{d}^{(2)}(K, N) &= (d(\mathbb{Z}_{K+1}^\infty, N+1) \ d(\mathbb{Z}_{K+1}^\infty, N+2) \ \cdots \ d(\mathbb{Z}_{K+1}^\infty, K)) \end{aligned}$$

定義より、 $\mathcal{E}(K, N) \in \mathbb{Z}_0^N$ は $\mathbb{Z}_{K+1}^\infty$ 内の各状態 $i \in \mathbb{Z}_{K+1}^\infty$ から $\mathbb{Z}_0^N$ へ初到達した際、マルコフ連鎖が取りうる状態の集合 $\mathcal{E}_i(K, N)$ の和集合 $\cup_{i \in \mathbb{Z}_{K+1}^\infty} \mathcal{E}_i(K, N)$ と同一である。

行列 $\mathbf{H}(K, N)$ ならびに $\mathbf{H}(K, N)$ の各行の和が 1 となるように正規化した行列 $\overline{\mathbf{H}}(K, N)$ を次式で定義する。

$$\mathbf{H}(K, N) = \begin{cases} (-\mathbf{Q}^{(1,1)}(N, N))^{-1}, & K = N \\ (-\mathbf{Q}^{(1,1)}(K, N) - \mathbf{Q}^{(1,2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1}\mathbf{Q}^{(2,1)}(K, N))^{-1}, & K > N \end{cases} \quad (8)$$

$$\overline{\mathbf{H}}(K, N) = \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})\mathbf{H}(K, N) = \begin{pmatrix} \bar{h}_0(K, N) \\ \bar{h}_1(K, N) \\ \vdots \\ \bar{h}_N(K, N) \end{pmatrix} \quad (9)$$

ただし、任意のベクトル $\mathbf{x}$ に対して $\text{diag}(\mathbf{x})$ は $[\mathbf{x}]_i$ を i 番目の対角要素にもつ対角行列を表す。定義より、 $\mathbf{Q}^{(2,2)}(K, N)$ は正則である。加えて、 $\mathbf{Q}^{(1,1)}(K, N) + \mathbf{Q}^{(1,2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1}\mathbf{Q}^{(2,1)}(K, N)$ も正則であるため、 $\mathbf{H}(K, N)$ は矛盾なく定義され、 $\mathbf{H}(K, N) \geq \mathbf{0}$ を満たす。さらに $\mathbf{H}(K, N)\mathbf{e} > \mathbf{0}$ であるため、 $\overline{\mathbf{H}}(K, N)$ も矛盾なく定義される。式 (9) より、 $\overline{\mathbf{H}}(K, N)$ は正則な確率行列である。よって、 $\bar{h}_i(K, N)$ ($i \in \mathbb{Z}_0^N$) は確率ベクトルであり、かつ、 $(N+1)$ 次元空間上の基底ベクトルであることに注意する。

$\mathcal{P}^+(K, N)$ を $\bar{h}_i(K, N)$ ($i \in \mathcal{E}(K, N)$) で張られる凸多面体と定義する。

$$\mathcal{P}^+(K, N) = \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \sum_{i \in \mathcal{E}(K, N)} \alpha_i \bar{h}_i(K, N), \alpha_i \geq 0 \ (i \in \mathcal{E}(K, N)), \sum_{i \in \mathcal{E}(K, N)} \alpha_i = 1 \right\} \quad (10)$$

命題 1 ([7, 定理 5]). $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える。任意の $N, K \in \mathbb{Z}^+$ ($K \geq N$) に対して次式が成立する。

$$\pi(N) \in \text{ri } \mathcal{P}^+(K, N)$$

ただし、 $\text{ri } \mathcal{P}^+(K, N)$ は凸多面体 $\mathcal{P}^+(K, N)$ の相対的内部を表す。

$$\text{ri } \mathcal{P}^+(K, N) = \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \sum_{i \in \mathcal{E}(K, N)} \alpha_i \bar{h}_i(K, N), \alpha_i > 0 \ (i \in \mathcal{E}(K, N)), \sum_{i \in \mathcal{E}(K, N)} \alpha_i = 1 \right\} \quad (11)$$

$\bar{h}_i(K, N)$ ($i \in \mathbb{Z}_0^N$) は線形独立であるので、 $\mathcal{P}^+(K, N)$ は $N+1$ 次元空間上の $|\mathcal{E}(K, N)| - 1$ -単体である。これらの単体に関して次の性質が示されている。

命題 2 ([7, 定理 7,8]). $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ において以下が成立する。

$$\mathcal{P}^+(K_2, N) \subseteq \mathcal{P}^+(K_1, N) \quad (N \leq K_1 < K_2), \quad \bigcap_{k=N}^{\infty} \mathcal{P}^+(k, N) = \lim_{k \rightarrow \infty} \mathcal{P}^+(k, N) = \{\pi(N)\}$$

3. 線形不等式系による $\pi(N)$ の特徴づけ

この節では、命題 1、ならびに、命題 2 で示された包含関係を線形不等式系という観点から再考する。その準備として、 $(N+1)$ 次元空間上の凸多面体に関する既知の結果をまとめておく。まず、適当な行列 $\mathbf{A}$, $\mathbf{B}$ を用いて、 $(N+1)$ 次元空間上の凸多面体 $\mathcal{P}$ を次式で定義する。

$$\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}\mathbf{A} \geq \mathbf{0}, \mathbf{x}\mathbf{B} = \mathbf{0}, \mathbf{x}\mathbf{e} = 1\} \quad (12)$$

一般に、上記のような凸多面体 $\mathcal{P}$ は有限生成凸錐と超平面 $\{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}\mathbf{e} = 1\}$ の共通部分で表現可能である [6, Theorem 2.15]. すなわち、 $\mathcal{P}$ は $\overline{\mathbf{C}}\mathbf{e} = \mathbf{e}$ を満たす適当な行列 $\overline{\mathbf{C}}$ を用いて以下のようにも表現できる.

$$\begin{aligned}\mathcal{P} &= \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \overline{\mathbf{C}}, \alpha \geq 0\} \cap \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}\mathbf{e} = 1\} \\ &= \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \overline{\mathbf{C}}, \alpha \geq 0, \alpha \mathbf{e} = 1\}\end{aligned}\quad (13)$$

補題 3. 式 (12) における $\mathbf{A}$, $\mathbf{B}$ がそれぞれ $(N+1) \times (M+1)$ 行列, $(N+1) \times (N-M)$ 行列であり, 行列 $(\mathbf{A} \ \mathbf{B})$ が正則であるとき, 式 (13) における $\overline{\mathbf{C}}$ は $(M+1) \times (N+1)$ 行列であり, その (i, j) 要素は次式で与えられる.

$$[\overline{\mathbf{C}}]_{i,j} = \frac{[(\mathbf{A} \ \mathbf{B})^{-1}]_{i,j}}{[(\mathbf{A} \ \mathbf{B})^{-1}\mathbf{e}]_i}, \quad i \in \mathbb{Z}_0^M, j \in \mathbb{Z}_0^N \quad (14)$$

定義より, 行列 $\overline{\mathbf{C}}$ は逆行列 $(\mathbf{A} \ \mathbf{B})^{-1}$ における i 番目 ($i \in \mathbb{Z}_0^M$) の行を行和が 1 となるように正規化したベクトルで構成される. また, 凸多面体 $\mathcal{P}$ の相対的内部 $\text{ri } \mathcal{P}$ は次式で与えられる.

$$\begin{aligned}\text{ri } \mathcal{P} &= \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}\mathbf{A} > \mathbf{0}, \mathbf{x}\mathbf{B} = \mathbf{0}, \mathbf{x}\mathbf{e} = 1\} \\ &= \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \overline{\mathbf{C}}, \alpha > 0, \alpha \mathbf{e} = 1\}\end{aligned}$$

以下では, 最初に $K = N$ の場合について議論した後, $K > N$ の場合について考察する.

3.1. $K = N$ の場合

$K = N$ の場合, すなわち, 式 (2) で与えた分割を考える. 式 (1) および 式 (2) より

$$\pi^{(1)}(N)\mathbf{Q}^{(1,1)}(N) + \pi^{(2)}(N)\mathbf{Q}^{(2,1)}(N) = \mathbf{0}$$

が成立する. よって, 式 (3) より, 条件付き定常分布 $\pi(N)$ は次式を満たす.

$$\pi(N)(-\mathbf{Q}^{(1,1)}(N)) = \frac{\pi^{(2)}(N)}{\pi^{(1)}(N)\mathbf{e}} \cdot \mathbf{Q}^{(2,1)}(N) \quad (15)$$

ここで, $\pi^{(2)}(N)/\pi^{(1)}(N)\mathbf{e} > \mathbf{0}$ ならびに $\mathbf{Q}^{(2,1)}(N) \geq \mathbf{0}$ であるので, 式 (15) の右辺は非負である. よって

$$\pi(N)(-\mathbf{Q}^{(1,1)}(N)) \geq \mathbf{0}$$

を得る. そこで凸多面体 $\mathcal{P}(N)$ を次式で定義する.

$$\mathcal{P}(N) = \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}(-\mathbf{Q}^{(1,1)}(N)) \geq \mathbf{0}, \mathbf{x}\mathbf{e} = 1\} \quad (16)$$

明らかに $\pi(N) \in \mathcal{P}(N)$ である.

式 (16) で定義された凸多面体 $\mathcal{P}(N)$ の別表現を与えるため, 行列 $\mathbf{H}(N) := \mathbf{H}(N, N)$ ならびに $\mathbf{H}(N)$ の各行の和が 1 となるように正規化した行列 $\overline{\mathbf{H}}(N) := \overline{\mathbf{H}}(N, N)$ を定義する.

$$\mathbf{H}(N) = (-\mathbf{Q}^{(1,1)}(N))^{-1}, \quad \overline{\mathbf{H}}(N) = \text{diag}^{-1}(\mathbf{H}(N)\mathbf{e})\mathbf{H}(N) = \begin{pmatrix} \overline{h}_0(N) \\ \overline{h}_1(N) \\ \vdots \\ \overline{h}_N(N) \end{pmatrix} \quad (17)$$

上記の議論ならびに補題 3 より以下の定理を得る.

定理 4. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える. 任意の $N \in \mathbb{Z}^+$ に対して, 次式が成立する.

$$\pi(N) \in \mathcal{P}(N) \quad (18)$$

ただし, $\mathcal{P}(N)$ は式 (16) あるいは次式で与えられる.

$$\mathcal{P}(N) = \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \overline{\mathbf{H}}(N), \alpha \in \mathbb{R}^{N+1}, \alpha \geq \mathbf{0}, \alpha \mathbf{e} = 1\} \quad (19)$$

定理 4 より, 条件付き定常分布は適当な $(N+1)$ 次元確率ベクトル $\alpha^*(N)$ を用いて, $\pi(N) = \alpha^*(N)\overline{H}(N)$ と表現可能であることが分かる. ここで, 行列 $\overline{H}(N)$ は推移率行列 Q の北西角に位置する $Q^{(1,1)}(N)$ のみによって定まる $(N+1)$ 次元空間上の非負基底ベクトル $\overline{h}_i(N) (\geq \mathbf{0})$ で構成されていることに注意する. 一方, これらの基底ベクトルの重みを定める確率ベクトル $\alpha^*(N)$ は推移率行列 Q 全体によって定まる. なお, この事実は, 式 (15) ならびに 式 (17) から直ちに得ることができる.

$$\pi(N) = \frac{1}{\pi^{(1)}(N)e} \cdot \pi^{(2)}(N)Q^{(2,1)}(N)(-Q^{(1,1)}(N))^{-1} = \alpha^*(N)\overline{H}(N) \quad (20)$$

ただし $\alpha^*(N) = (\alpha_0^*(N) \alpha_1^*(N) \cdots \alpha_N^*(N))$ は次式で与えられる.

$$\alpha^*(N) = \frac{1}{\pi^{(1)}(N)e} \cdot \pi^{(2)}(N)Q^{(2,1)}(N)\text{diag}(\mathbf{H}(N)e) = \frac{\pi^{(2)}(N)Q^{(2,1)}(N)\text{diag}(\mathbf{H}(N)e)}{\pi^{(2)}(N)Q^{(2,1)}(N)\mathbf{H}(N)e}$$

$\overline{H}(N)$, $\alpha^*(N)$ はそれぞれ以下のような確率的意味を持つ. $\mathbf{H}(N)$ の (i, j) 要素は状態 $i \in \mathbb{Z}_0^N$ から状態集合 $\mathbb{Z}_{N+1}^\infty$ へ初到達するまでに状態 $j \in \mathbb{Z}_0^N$ に滞在する平均時間を表している. よって, $[\overline{H}(N)]_{i,j}$ ($i, j \in \mathbb{Z}_0^N$) は次式を満たす.

$$[\overline{H}(N)]_{i,j} = [\overline{h}_i(N)]_j = \frac{E(T_j(Z_{N+1}^\infty) \mid X(0) = i)}{\sum_{m \in \mathbb{Z}_0^N} E(T_m(Z_{N+1}^\infty) \mid X(0) = i)}, \quad i, j \in \mathbb{Z}_0^N \quad (21)$$

ただし, $T_j(Z_{N+1}^\infty)$ ($j \in \mathbb{Z}_0^N$) は $\mathbb{Z}_{N+1}^\infty$ へ初到達するまでに状態 j に滞在する延べ時間である.

τ を $X(\tau) \in \mathbb{Z}_0^N$ であるような時点とする. $\{X(t); t \geq 0\}$ は定常なので, $\Pr(X(\tau) = j) = \Pr(X(\tau) = j \mid X(\tau) \in \mathbb{Z}_0^N) = \pi_j(N)$ を満たす. ここで, $\mathbb{Z}_{N+1}^\infty$ 内の状態から $\mathbb{Z}_0^N$ 内の状態への遷移が発生した直後から, 再び $\mathbb{Z}_{N+1}^\infty$ 内の状態へ遷移するまでの期間に着目する. 以下では, この期間を $\mathbb{Z}_0^N$ 滞在期間と呼ぶ. このとき, $\alpha_i^*(N)$ は時点 τ を含む $\mathbb{Z}_0^N$ 滞在期間が状態 i から開始される確率と解釈される. 一方, $[\overline{h}_i(N)]_j$ は時点 τ が状態 i から開始した $\mathbb{Z}_0^N$ 滞在期間に含まれているという条件下で, $X(\tau) = j$ である確率と解釈される. すなわち, 式 (20) は定常状態確率 $\pi_j(N)$ を $\mathbb{Z}_0^N$ 滞在期間の開始時点における状態で場合分けした表現と見ることができる.

次に, 式 (2) の分割において推移率行列の北西角 $Q^{(1,1)}(N)$ の直下にある $Q^{(2,1)}(N)$ の部分情報を考慮した場合を考える. まず, 状態集合 $\mathcal{E}(i)$ ($i \in \mathbb{Z}^+$) を次式で定義する.

$$\mathcal{E}(i) = \mathcal{E}(i, i) = \{j \in \mathbb{Z}_0^i; d(\mathbb{Z}_{i+1}^\infty, j) = 1\}, \quad i \in \mathbb{Z}^+ \quad (22)$$

よって, $\mathcal{E}(N)$ は $N+1$ 以上のいずれかの状態から直接遷移可能な N 以下の状態の集合を表す. エルゴード的であるという仮定より $\mathcal{E}(N) \neq \emptyset$ である. 以下では, 一般性を失うことなく, $\mathcal{E}(N) = \{0, 1, \dots, |\mathcal{E}(N)| - 1\}$ と仮定する.

$\mathcal{E}(N) \neq \mathbb{Z}_0^N$ の場合を考える. $\mathcal{E}(N)$ に従って $Q^{(1,1)}(N)$ ならびに $Q^{(2,1)}(N)$ をそれぞれ二つの部分行列に分割する.

$$\begin{pmatrix} Q^{(1,1)}(N) \\ Q^{(2,1)}(N) \end{pmatrix} = \begin{matrix} \mathbb{Z}_0^N & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \mathbb{Z}_{N+1}^\infty & \mathbb{Z}_0^N \end{matrix} \begin{pmatrix} Q_+^{(1,1)}(N) & Q_0^{(1,1)}(N) \\ Q_+^{(2,1)}(N) & \mathbf{0} \end{pmatrix}$$

このとき, 式 (15) より次式を得る.

$$\pi(N)(-Q_+^{(1,1)}(N) - Q_0^{(1,1)}(N)) = \begin{pmatrix} \mathcal{E}(N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \frac{\pi^{(2)}(N)}{\pi^{(1)}(N)e} \cdot Q_+^{(2,1)}(N) & \mathbf{0} \end{pmatrix}$$

さらに $\pi^{(2)}(N)Q_+^{(2,1)}(N) > \mathbf{0}$ に注意し, 補題 3 を用いると以下の定理を得る.

定理 5. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える. もし, $\mathcal{E}(N) \neq \mathbb{Z}_0^N$ ならば, 任意の $N \in \mathbb{Z}^+$ に対して, 次式が成立する.

$$\pi(N) \in \text{ri } \mathcal{P}^+(N)$$

ただし

$$\begin{aligned} \mathcal{P}^+(N) &= \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}(-Q_+^{(1,1)}(N)) \geq \mathbf{0}, \mathbf{x}(-Q_0^{(1,1)}(N)) = \mathbf{0}, \mathbf{x}e = 1 \right\} \\ &= \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \sum_{i \in \mathcal{E}(N)} \alpha_i \overline{h}_i(N), \alpha_i \geq 0 (i \in \mathcal{E}(N)), \sum_{i \in \mathcal{E}(N)} \alpha_i = 1 \right\} \end{aligned} \quad (23)$$

定義より, $\mathcal{P}^+(N) \subset \mathcal{P}(N)$ であることに注意する. 定理 5 は推移率行列の北西角 $\mathbf{Q}^{(1,1)}(N)$ の直下にある $\mathbf{Q}^{(2,1)}(N)$ が持つ部分的な情報を利用して, $\mathbb{Z}_{N+1}^\infty$ から直接遷移できない $\mathbb{Z}_0^N$ 内の状態を場合分けにおいて除外した結果であり, $\pi(N)$ が $\bar{\mathbf{h}}_i(N)$ ($i \in \mathcal{E}(N)$) を正の確率で混合したもので構成されることを示している.

$$\pi(N) = \sum_{i \in \mathcal{E}(N)} \alpha_i^*(N) \bar{\mathbf{h}}_i(N), \quad \alpha_i^*(N) > 0 \ (i \in \mathcal{E}(N)), \quad \sum_{i \in \mathcal{E}(N)} \alpha_i^*(N) = 1 \quad (24)$$

すなわち $\pi(N)$ の求解は式 (24) の制約を満たす $\{\alpha_i^*(N); i \in \mathcal{E}(N)\}$ の求解と等価である.

3.2. $K > N$ の場合

$K > N$ の場合, すなわち, 式 (5) で与えた分割を考える. 式 (1), 式 (4) および式 (5) より,

$$\begin{aligned} \pi^{(1)}(K, N) \mathbf{Q}^{(1,1)}(K, N) + \pi^{(2)}(K, N) \mathbf{Q}^{(2,1)}(K, N) + \pi^{(3)}(K, N) \mathbf{Q}^{(3,1)}(K, N) &= \mathbf{0} \\ \pi^{(1)}(K, N) \mathbf{Q}^{(1,2)}(K, N) + \pi^{(2)}(K, N) \mathbf{Q}^{(2,2)}(K, N) + \pi^{(3)}(K, N) \mathbf{Q}^{(3,2)}(K, N) &= \mathbf{0} \end{aligned}$$

が成立する. これらから $\mathbf{Q}^{(2,2)}(K, N)$ が正則であることを利用して $\pi^{(2)}(K, N)$ を消去し, 式 (3) を用いると, 以下のような式 (15) と同型の結果が得られる.

$$\pi(N)(-\tilde{\mathbf{Q}}^{(1,1)}(K, N)) = \frac{\pi^{(3)}(K, N)}{\pi^{(1)}(K, N)e} \cdot \tilde{\mathbf{Q}}^{(3,1)}(K, N) \quad (25)$$

ただし

$$\begin{aligned} \tilde{\mathbf{Q}}^{(1,1)}(K, N) &= \mathbf{Q}^{(1,1)}(K, N) + \mathbf{Q}^{(1,2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1} \mathbf{Q}^{(2,1)}(K, N) \\ \tilde{\mathbf{Q}}^{(3,1)}(K, N) &= \mathbf{Q}^{(3,1)}(K, N) + \mathbf{Q}^{(3,2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1} \mathbf{Q}^{(2,1)}(K, N) \end{aligned} \quad (26)$$

なお, この操作は, 中間に位置する状態集合 $\mathbb{Z}_{N+1}^K$ をセンサーアウトして, 状態集合 $\mathbb{Z}_0^N \cup \mathbb{Z}_{K+1}^\infty$ 上で定義された以下の推移率行列 $\tilde{\mathbf{Q}}(K, N)$ をもつマルコフ連鎖を考えることと等価であることに注意する.

$$\tilde{\mathbf{Q}}(K, N) = \begin{matrix} & \mathbb{Z}_0^N & \mathbb{Z}_{K+1}^\infty \\ \begin{matrix} \mathbb{Z}_0^N \\ \mathbb{Z}_{K+1}^\infty \end{matrix} & \begin{pmatrix} \tilde{\mathbf{Q}}^{(1,1)}(K, N) & \tilde{\mathbf{Q}}^{(1,3)}(K, N) \\ \tilde{\mathbf{Q}}^{(3,1)}(K, N) & \tilde{\mathbf{Q}}^{(3,3)}(K, N) \end{pmatrix} \end{matrix} \quad (27)$$

ただし,

$$\tilde{\mathbf{Q}}^{(i,3)}(K, N) = \mathbf{Q}^{(i,3)}(K, N) + \mathbf{Q}^{(i,2)}(K, N)(-\mathbf{Q}^{(2,2)}(K, N))^{-1} \mathbf{Q}^{(2,3)}(K, N), \quad i = 1, 3$$

すなわち, 式 (27) に基づけば, $K = N$ の場合と同じ議論が $K > N$ の場合にも適用可能である. 具体的には, 式 (8) ならびに式 (9) で定義したように

$$\mathbf{H}(K, N) = (-\tilde{\mathbf{Q}}^{(1,1)}(K, N))^{-1}, \quad \bar{\mathbf{H}}(K, N) = \text{diag}^{-1}(\mathbf{H}(K, N)e)\mathbf{H}(K, N)$$

として, 式 (25) を元に $K = N$ の場合と同様の議論を行うと, $K = N$ の場合に対する定理 4 に対応する以下の結果を得ることができる.

定理 6. $\mathbb{Z}^+$ 上で定義されるエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える. 任意の $N, K \in \mathbb{Z}^+$ ($K > N$) に対して次式が成立する.

$$\pi(N) \in \mathcal{P}(K, N) \quad (28)$$

ただし, $\mathcal{P}(K, N)$ は次式で与えられる.

$$\mathcal{P}(K, N) = \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}(-\tilde{\mathbf{Q}}^{(1,1)}(K, N)) \geq \mathbf{0}, \mathbf{x}e = 1 \right\} \quad (29)$$

$$= \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \bar{\mathbf{H}}(K, N), \alpha \in \mathbb{R}^{N+1}, \alpha \geq \mathbf{0}, \alpha e = 1 \right\} \quad (30)$$

さらに式 (7) で定義された $\mathcal{E}(K, N)$ を用いて定理 5 と同様の議論を行うことにより, 命題 1 を得ることができる. なお, 命題 1 の式 (11) のもととなっている式 (10) で与えられる凸多面体 $\mathcal{P}^+(K, N)$ は

$$\tilde{\mathbf{Q}}^{(1,1)}(K, N) = \begin{pmatrix} \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(K, N) \\ \tilde{\mathbf{Q}}_+^{(1,1)}(K, N) & \tilde{\mathbf{Q}}_0^{(1,1)}(K, N) \end{pmatrix}$$

としたとき、補題 3 より、以下のような別表現をもつことが分かる。

$$\mathcal{P}^+(K, N) = \begin{cases} \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}(-\tilde{\mathbf{Q}}^{(1,1)}(K, N)) \geq \mathbf{0}, \mathbf{x}\mathbf{e} = 1 \right\}, & \mathcal{E}(K, N) = \mathbb{Z}_0^N \\ \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x}(-\tilde{\mathbf{Q}}_+^{(1,1)}(K, N)) \geq \mathbf{0}, \mathbf{x}(-\tilde{\mathbf{Q}}_0^{(1,1)}(K, N)) = \mathbf{0}, \mathbf{x}\mathbf{e} = 1 \right\}, & \mathcal{E}(K, N) \neq \mathbb{Z}_0^N \end{cases}$$

定義より、 $\mathcal{E}(K, N) = \mathbb{Z}_0^N$ ならば $\mathcal{P}^+(K, N) = \mathcal{P}(K, N)$ 、 $\mathcal{E}(K, N) \neq \mathbb{Z}_0^N$ ならば $\mathcal{P}^+(K, N) \subset \mathcal{P}(K, N)$ であることを注意する。

3.3. 凸多面体間の包含関係

本節では、命題 2 で示されている凸多面体間の包含関係を議論する。最初に式 (16) と式 (29) を考える。その準備として、式 (26) で与えた $\tilde{\mathbf{Q}}^{(1,1)}(K, N)$ を $\mathbf{Q}^{(1,1)}(K, N) = \mathbf{Q}^{(1,1)}(N)$ を用いて以下のように書き換える。

$$\tilde{\mathbf{Q}}^{(1,1)}(K, N) = \mathbf{Q}^{(1,1)}(N)(\mathbf{I} - \mathbf{R}(K, N)) \quad (31)$$

ただし

$$\mathbf{R}(K, N) = (-\mathbf{Q}^{(1,1)}(K, N))^{-1} \mathbf{Q}^{(1,2)}(K, N) \cdot (-\mathbf{Q}^{(2,2)}(K, N))^{-1} \mathbf{Q}^{(2,1)}(K, N)$$

(31) において $\tilde{\mathbf{Q}}^{(1,1)}(K, N)$ ならびに $\mathbf{Q}^{(1,1)}(N)$ は共に正則であるので、 $(\mathbf{I} - \mathbf{R}(K, N))$ も正則である。また、 $\mathbf{R}(K, N)$ の確率的な解釈より、 $(\mathbf{I} - \mathbf{R}(K, N))^{-1}$ は非負行列となることが示される。よって $\mathbf{x}\tilde{\mathbf{Q}}^{(1,1)}(K, N) \geq \mathbf{0}$ の両辺に非負行列 $(\mathbf{I} - \mathbf{R}(K, N))^{-1}$ を掛けると $\mathbf{x}\mathbf{Q}^{(1,1)}(N) \geq \mathbf{0}$ が得られる。

$$\mathbf{x}\tilde{\mathbf{Q}}^{(1,1)}(K, N) \geq \mathbf{0} \Rightarrow \mathbf{x}\mathbf{Q}^{(1,1)}(N) \geq \mathbf{0}$$

すなわち、式 (16)、式 (29) より

$$\mathcal{P}(K, N) \subseteq \mathcal{P}(N) \quad (32)$$

が成立することがわかる。

次に、式 (30) で与えられる凸多面体の別表現を考える。式 (31) より

$$\begin{aligned} \mathbf{H}(K, N) &= (-\tilde{\mathbf{Q}}^{(1,1)}(K, N))^{-1} = (\mathbf{I} - \mathbf{R}(K, N))^{-1} (-\mathbf{Q}^{(1,1)}(N))^{-1} \\ &= (\mathbf{I} - \mathbf{R}(K, N))^{-1} \mathbf{H}(N) \end{aligned}$$

を得る。よって、式 (9) に注意すると、 $\beta \in \mathbb{Z}_0^N$ に対して

$$\begin{aligned} \beta \bar{\mathbf{H}}(K, N) &= \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1} \mathbf{H}(N) \\ &= \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1} \text{diag}(\mathbf{H}(N)\mathbf{e}) \bar{\mathbf{H}}(N) \\ &= \alpha \bar{\mathbf{H}}(N) \end{aligned} \quad (33)$$

を得る。ただし

$$\alpha = \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1} \text{diag}(\mathbf{H}(N)\mathbf{e}) \quad (34)$$

式 (34) において、 $\text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})$ 、 $(\mathbf{I} - \mathbf{R}(K, N))^{-1}$ 、 $\text{diag}(\mathbf{H}(N)\mathbf{e})$ は全て非負行列である。よって、 $\beta \geq \mathbf{0}$ ならば $\alpha \geq \mathbf{0}$ が成立する。また、

$$\alpha \mathbf{e} = \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1} \mathbf{H}(N)\mathbf{e} = \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e}) \mathbf{H}(K, N)\mathbf{e} = \beta \mathbf{e}$$

である。これらは、 β が確率ベクトルであれば α も確率ベクトルとなることを示している。

ここで

$$\begin{aligned} \Gamma(N) &= \left\{ \alpha \in \mathbb{R}^{N+1}; \alpha \geq \mathbf{0}, \alpha \mathbf{e} = 1 \right\} \\ \Gamma(K, N) &= \left\{ \alpha \in \mathbb{R}^{N+1}; \alpha = \beta \text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1} \text{diag}(\mathbf{H}(N)\mathbf{e}), \beta \in \Gamma(N) \right\} \end{aligned}$$

とする。定義より

$$\Gamma(K, N) \subset \Gamma(N) \quad (35)$$

が成立する。さらに、式 (19)、式 (30) はこれらを用いて以下のように書き換えることができる。

$$\mathcal{P}(N) = \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \bar{\mathbf{H}}(N), \alpha \in \Gamma(N) \right\}, \quad \mathcal{P}(K, N) = \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \alpha \bar{\mathbf{H}}(N), \alpha \in \Gamma(K, N) \right\}$$

よって、式 (35) に注意すると式 (32) が成立することが分かる。

次に $\mathcal{P}^+(N)$ と $\mathcal{P}^+(K, N)$ の包含関係を議論する。 $\mathcal{E}(K, N) \subseteq \mathcal{E}(N)$ なので、必要に応じて $\mathbb{Z}_0^N$ を $\mathcal{E}(K, N)$, $\mathcal{E}(N) \setminus \mathcal{E}(K, N)$, $\mathbb{Z}_0^N \setminus \mathcal{E}(N)$ の三つに分割する。例えば

$$\mathbf{Q}^{(2,1)}(K, N) = \begin{pmatrix} \mathcal{E}(K, N) & \mathcal{E}(N) \setminus \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \mathbf{Q}_+^{(2,1)}(K, N) & \mathbf{Q}_{++}^{(2,1)}(K, N) & \mathbf{O} \end{pmatrix} \quad (36)$$

また、 $(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + (\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}$ より、ある非負行列 $\mathbf{B}$ を用いて

$$(\mathbf{I} - \mathbf{R}(K, N))^{-1} = \mathbf{I} + \mathbf{B}\mathbf{Q}^{(2,1)}(K, N), \quad \mathbf{B} \geq \mathbf{O} \quad (37)$$

と書くことができる。そこで $\mathbf{x} \in \mathcal{P}^+(K, N)$, すなわち、

$$\mathbf{x}(-\tilde{\mathbf{Q}}^{(1,1)}(K, N)) = \begin{pmatrix} \mathcal{E}(K, N) & \mathcal{E}(N) \setminus \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \mathbf{y} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{y} \geq \mathbf{0}$$

ならば、式 (31), 式 (36), 式 (37) より

$$\begin{aligned} \mathbf{x}(-\mathbf{Q}^{(1,1)}(N)) &= (\mathbf{y} \ \mathbf{0} \ \mathbf{0})(\mathbf{I} - \mathbf{R}(K, N))^{-1} = (\mathbf{y} \ \mathbf{0} \ \mathbf{0}) + (\mathbf{y} \ \mathbf{0} \ \mathbf{0})\mathbf{B}(\mathbf{Q}_+^{(2,1)}(K, N) \ \mathbf{Q}_{++}^{(2,1)}(K, N) \ \mathbf{O}) \\ &= (\mathbf{y}_1 \ \mathbf{y}_2 \ \mathbf{0}), \quad \mathbf{y}_1, \mathbf{y}_2 \geq \mathbf{0} \end{aligned}$$

を得る。すなわち

$$\mathcal{P}^+(K, N) \subseteq \mathcal{P}^+(N) \quad (38)$$

が成立する。

一方、式 (33) において

$$\boldsymbol{\beta} = \begin{pmatrix} \mathcal{E}(K, N) & \mathcal{E}(N) \setminus \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \boldsymbol{\beta}_1 & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\beta}_1 \geq \mathbf{0}, \boldsymbol{\beta}_1 \mathbf{e} = 1$$

とおくと、式 (34) より

$$\boldsymbol{\alpha} = (\boldsymbol{\beta}_1 \ \mathbf{0} \ \mathbf{0})\text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1}\text{diag}(\mathbf{H}(N)\mathbf{e})$$

となり、式 (36) と式 (37) に注意すると $\boldsymbol{\alpha}$ は以下のように書くことができる。

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathcal{E}(K, N) & \mathcal{E}(N) \setminus \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \boldsymbol{\alpha}'_1 & \boldsymbol{\alpha}'_2 & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\alpha}'_1, \boldsymbol{\alpha}'_2 \geq \mathbf{0}, \boldsymbol{\alpha}'_1 \mathbf{e} + \boldsymbol{\alpha}'_2 \mathbf{e} = 1$$

よって

$$\begin{aligned} \Gamma^+(N) &= \left\{ \boldsymbol{\alpha} = \begin{pmatrix} \mathcal{E}(N) & \mathbb{Z}_0^N \setminus \mathcal{E}(N) \\ \boldsymbol{\alpha}_1 & \mathbf{0} \end{pmatrix} \in \mathbb{R}^{N+1}; \boldsymbol{\alpha}_1 \geq \mathbf{0}, \boldsymbol{\alpha}_1 \mathbf{e} = 1 \right\} \\ \Gamma^+(K, N) &= \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{N+1}; \boldsymbol{\alpha} = \boldsymbol{\beta}\text{diag}^{-1}(\mathbf{H}(K, N)\mathbf{e})(\mathbf{I} - \mathbf{R}(K, N))^{-1}\text{diag}(\mathbf{H}(N)\mathbf{e}), \right. \\ &\quad \left. \boldsymbol{\beta} = \begin{pmatrix} \mathcal{E}(K, N) & \mathbb{Z}_0^N \setminus \mathcal{E}(K, N) \\ \boldsymbol{\beta}_1 & \mathbf{0} \end{pmatrix}, \boldsymbol{\beta}_1 \geq \mathbf{0}, \boldsymbol{\beta}_1 \mathbf{e} = 1 \right\} \end{aligned}$$

とすると $\Gamma^+(K, N) \subset \Gamma^+(N)$ が成立する。さらに、式 (23), 式 (10) はこれらを用いて以下のように書き換えることができる。

$$\begin{aligned} \mathcal{P}^+(N) &= \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \boldsymbol{\alpha}\overline{\mathbf{H}}(N), \boldsymbol{\alpha} \in \Gamma^+(N) \right\} \\ \mathcal{P}^+(K, N) &= \left\{ \mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \boldsymbol{\alpha}\overline{\mathbf{H}}(N), \boldsymbol{\alpha} \in \Gamma^+(K, N) \right\} \end{aligned}$$

よって、式 (38) が成立することが分かる。

ここまでは $K = N$ の場合と $K > N$ の場合の包含関係を議論してきたが、 $K = K_1$ と $K = K_2$ ($K_1 < K_2$) の場合の包含関係を示す命題 2 は、 $K = K_1$ の場合を式 (27) に基づいて $K = N$ の場合と同様に扱い、 $K = K_2$ の場合に対しては $K > N$ の場合に行った議論を適用することで示すことができる。

4. $K > N$ の場合における最小凸多面体

3.1 節で述べたとおり、条件付き定常状態確率 $\pi_j(N) = \Pr(X(\tau) = j \mid X(\tau) \in \mathbb{Z}_0^N)$ は時刻 τ が属する $\mathbb{Z}_0^N$ 滞在期間の開始時点における状態 $i \in \mathbb{Z}_0^N$ による場合分けで表現することができる。本節では $K > N$ の場合を想定し、 $\mathbb{Z}_{K+1}^\infty$ 内の状態から $\mathbb{Z}_0^N$ への初到達において、状態集合 $\mathbb{Z}_{N+1}^K$ 内の各状態への訪問に関する情報を利用して場合分けを行うことを考える。これにより、推移率行列の北西角に位置する $(K+1) \times (K+1)$ 部分行列が条件付き定常分布 $\pi(N)$ に関して保持している最も詳細な情報、すなわち、式 (10) で与えられる単体より小さな、 $\pi(N)$ を相対的内点にもつ極小の凸多面体が同定される。

まず初めに、状態集合 $\mathbb{Z}^+$ の任意の部分集合 $\mathcal{B}$ に対して、 $F(\mathcal{B})$ を $\mathcal{B}$ への初到達時間と定義する。

$$F(\mathcal{B}) = \inf\{t \geq 0; X(t) \in \mathcal{B}\}, \quad \mathcal{B} \subset \mathbb{Z}^+$$

定義 7 ((K, N) -skip-free 集合). $\mathbb{Z}_0^K$ の部分集合 $\mathcal{X}$ が以下の条件を満たすとき、 $\mathcal{X}$ を (K, N) -skip-free 集合と呼ぶ。

$$\Pr(X(t) \in \mathcal{X} \text{ for some } t \in (0, F(\mathbb{Z}_0^N)] \mid X(0) \in \mathbb{Z}_{K+1}^\infty) = 1 \quad (39)$$

かつ、 $\mathcal{X}$ の任意の真部分集合 $\mathcal{Y}$ に対して

$$\Pr(X(t) \in \mathcal{Y} \text{ for some } t \in (0, F(\mathbb{Z}_0^N)] \mid X(0) \in \mathbb{Z}_{K+1}^\infty) < 1 \quad (40)$$

が成立する。

全ての (K, N) -skip-free 集合を要素に持つ集合族を $\mathcal{S}(K, N)$ で表す。定義より、 $\mathcal{X} \in \mathcal{S}(K, N)$ は $\mathbb{Z}_0^N$ への初到達時間が $\mathbb{Z}_{K+1}^\infty$ 内のいずれの状態から始まろうとも、 $\mathbb{Z}_0^N$ への初到達時間までに必ず一度は訪問される極小の状態集合である。言い換えると、 $\mathbb{Z}_0^K$ 滞在期間の開始時点から $\mathbb{Z}_0^N$ への初到達時点の間に必ず訪問される極小の状態集合である。なお、定義より、 $\mathcal{E}(K, N)$ は (K, N) -skip-free 集合であることに注意する。

このとき、 $\pi_j(N) = \Pr(X(\tau) = j \mid X(\tau) \in \mathbb{Z}_0^N)$ ($j \in \mathbb{Z}_0^N$) に関して、時刻 τ が属する $\mathbb{Z}_0^K$ 滞在期間において初めて skip-free 集合 $\mathcal{X}$ に遷移した時点における状態の場合分けを行うことを考える。その準備として、状態 $i \in \mathbb{Z}_{N+1}^K$ に対しても式 (21) と同様に $\bar{h}_i(K, N)$ を定義する。

$$[\bar{h}_i(K, N)]_j = \frac{\mathbb{E}(T_j(\mathbb{Z}_{K+1}^\infty) \mid X(0) = i)}{\sum_{m \in \mathbb{Z}_0^N} \mathbb{E}(T_m(\mathbb{Z}_{K+1}^\infty) \mid X(0) = i)}, \quad i \in \mathbb{Z}_{N+1}^K, j \in \mathbb{Z}_0^N.$$

なお、次式が成立する。

$$\begin{pmatrix} \bar{h}_0(K, N) \\ \bar{h}_1(K, N) \\ \vdots \\ \bar{h}_K(K, N) \end{pmatrix} = \text{diag}(\mathbf{c}) \begin{pmatrix} \mathbf{I} \\ (-\mathbf{Q}^{(2,2)}(K, N))^{-1} \mathbf{Q}^{(2,1)}(K, N) \end{pmatrix} \mathbf{H}(K, N) \quad (41)$$

ただし、 $\mathbf{c}$ の各要素は $\bar{h}_i(K, N)\mathbf{e} = 1$ ($i \in \mathbb{Z}_0^K$) となるように定められる正規化定数である。

定理 8. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える。任意の $N, K \in \mathbb{Z}^+$ ($K > N$) に対して、次式が成立する。

$$\pi(N) \in \text{ri } \mathcal{P}(K, N; \mathcal{X}), \quad \mathcal{X} \in \mathcal{S}(K, N)$$

ただし、 $\mathcal{P}(K, N; \mathcal{X})$ は次式で定義される。

$$\mathcal{P}(K, N; \mathcal{X}) = \{\mathbf{x} \in \mathbb{R}^{N+1}; \mathbf{x} = \sum_{i \in \mathcal{X}} \alpha_i \bar{h}_i(K, N), \alpha_i \geq 0 (i \in \mathcal{X}), \sum_{i \in \mathcal{X}} \alpha_i = 1\} \quad (42)$$

定義より、 $\mathcal{P}^+(K, N) = \mathcal{P}(K, N; \mathcal{E}(K, N))$ であることに注意する。一般に、 N, K ($K > N$) の組に対して、 (K, N) -skip-free 集合は複数存在する。

定理 9. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える。任意の $N, K \in \mathbb{Z}^+$ ($K > N$) ならびに $\mathcal{X}_A, \mathcal{X}_B \in \mathcal{S}(K, N)$ に対して、次式が成立する。

$$\Pr(F(\mathcal{X}_A) \leq F(\mathcal{X}_B) \mid X(0) \in \mathbb{Z}_{K+1}^\infty) = 1 \quad \Rightarrow \quad \mathcal{P}(K, N; \mathcal{X}_A) \subseteq \mathcal{P}(K, N; \mathcal{X}_B)$$

定理 9 は常に $F(\mathcal{X}_A) \leq F(\mathcal{X}_B)$ であるような二つの skip-free 集合 $\mathcal{X}_A, \mathcal{X}_B$ が与えられたとき, $\mathbb{Z}_0^K$ 滞在期間内において, より早い時点で初到達可能な skip-free 集合 $\mathcal{X}_A$ を用いて構成された凸多面体 $\mathcal{P}(K, N; \mathcal{X}_A)$ に, $\pi(N)$ に関するより詳細な情報が含まれていることを示している. なお

$$\Pr(F(\mathcal{X}) \leq F(\mathcal{E}(K, N)) \mid X(0) \in \mathbb{Z}_{K+1}^\infty) = 1, \quad \mathcal{X} \in \mathcal{S}(K, N)$$

が任意の $\mathcal{X} \in \mathcal{S}(K, N)$ に対して成立する. すなわち, $\mathcal{E}(K, N)$ は極大の凸多面体を与える (K, N) -skip-free 集合である.

$\mathcal{D}(K, N)$ を次式を満たす skip-free 集合とする.

$$\mathcal{D}(K, N) \subseteq \mathcal{E}(K), \text{ かつ } \mathcal{D}(K, N) \in \mathcal{S}(K, N)$$

式 (22) より $\mathcal{E}(K)$ は $\mathbb{Z}_{K+1}^\infty$ のいずれかの状態から直接遷移可能な, K 以下の状態からなる集合である. よって, $\mathcal{D}(K, N)$ は $\mathcal{E}(K)$ 内の状態から式 (39) を満たすように不要な状態を取り除くことで得られる集合である. なお, $\mathcal{E}(K)$ は推移率行列の $(K+1) \times (K+1)$ 北西角の直下に位置する行列から得られる部分情報 $\{d(\mathbb{Z}_{K+1}^\infty, j) = 1; j \in \mathbb{Z}_0^K\}$ から求めることができる. さらに, この情報に加えて, 推移率行列の $(K+1) \times (K+1)$ 北西角がもつ情報を利用して, $\mathbb{Z}_0^N$ から $\mathcal{E}(K)$ へマルコフ連鎖の推移を逆向きに探索すれば, $O(K)$ の計算量で $\mathcal{D}(K, N)$ を同定できるが詳細は省略する.

定理 10. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える. 任意の $N, K \in \mathbb{Z}^+$ ($K > N$) に対して, $\mathcal{D}(K, N)$ がただ一つ存在し, 次式が成立する.

$$\Pr(F(\mathcal{D}(K, N)) \leq F(\mathcal{X}) \mid X(0) \in \mathbb{Z}_{K+1}^\infty) = 1, \quad \mathcal{X} \in \mathcal{S}(K, N)$$

定理 10 より, 次の系が直ちに導かれる.

系 11. $\mathbb{Z}^+$ 上で定義されたエルゴード的な連続時間マルコフ連鎖 $\{X(t); t \geq 0\}$ を考える. 任意の $N, K \in \mathbb{Z}^+$ ($K > N$) に対して, 次式が成立する.

$$\pi(N) \in \text{ri } \mathcal{P}(K, N; \mathcal{D}(K, N)) \subseteq \text{ri } \mathcal{P}(K, N; \mathcal{X}), \quad \mathcal{X} \in \mathcal{S}(K, N)$$

系 11 は, 推移率行列の北西角 $\mathbf{Q}^{(i,j)}(K, N)$ ($i, j = 1, 2$), ならびに, これらの直下に位置する部分行列が持つ推移に関する部分的な情報 $\{d(\mathbb{Z}_{K+1}^\infty, j); j \in \mathbb{Z}_0^K\}$ が与えられたとき, $\mathcal{P}(K, N; \mathcal{D}(K, N))$ が条件付き定常分布 $\pi(N)$ を特徴づける最小の凸多面体であることを示している.

5. おわりに

本稿では, 条件付き定常分布 $\pi(N)$ の線形不等式系による特徴づけを議論した. 2 節で示した文献 [7] の結果は, 元のマルコフ連鎖において $\mathbb{Z}_{K+1}^\infty$ をセンサーアウトした結果得られるマルコフ再生過程を考察することで導かれていた. 本稿では, これらの結果の大半が, 大域平衡方程式から直ちに得られる線形不等式系から自然に導かれることを示した. さらに, skip-free 集合という概念を導入し, 推移率行列の北西角 $\mathbf{Q}^{(i,j)}(K, N)$ ($i, j = 1, 2$), ならびに, これらの直下に位置する部分行列が持つ推移に関する部分的な情報 $\{d(\mathbb{Z}_{K+1}^\infty, j); j \in \mathbb{Z}_0^K\}$ が条件付き定常分布 $\pi(N)$ に関して保持している最も詳細な情報を示した.

参考文献

- [1] M. Kimura and T. Takine, Computing the conditional stationary distribution in Markov chains of level-dependent M/G/1-type, *Stochastic Models*, 34 (2018) 207–238.
- [2] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*; Society for Industrial and Applied Mathematics: Philadelphia, 1999.
- [3] M.F. Neuts, *Matrix-Geometric Solutions in Stochastic Models*, Johns Hopkins University Press, 1981.
- [4] M.F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Dekker, 1989.
- [5] T. Takine, Analysis and computation of the stationary distribution in a special class of level-dependent M/G/1-type and its application to BMAP/M/ ∞ and BMAP/M/c+M queues, *Queueing Systems*, 84 (2016) 49–77.
- [6] G.M. Ziegler, *Lectures on Polytopes*, Springer, 1995.
- [7] 木村雅俊, 滝根哲哉, レベル依存する G/G/1 型マルコフ連鎖における条件付き定常分布の誤差評価付き数値計算法, 2017 年度待ち行列シンポジウム報文集 (2018), 51–60.

大数の法則を用いたトーラス上のデータ拡散モデル

渡辺 樹 豊泉 洋
早稲田大学基幹理工学研究科

概要：生成・死滅を含む非線形反応拡散方程式による連続モデルとランダムウォークによるマルコフ過程モデルとを比較する。個体数に関して非線形な生成・死滅レートを持つマルコフモデルに対して異なるスケーリングで、位置の連続化と個体の流体化を行うことで非線形反応拡散方程式に L^∞ ノルムの意味で確率収束するような確率モデルの導入が可能であることを示す。

1. はじめに

1980年にL.Arnold, M.Theodosopuluにより化学反応拡散確率モデルが考えられ、普段我々が扱うような微分方程式で表される連続モデルとの比較の研究が進められた[1]。その後P.Kotelenez([7],[8],[9],[10],[11],[12])によって q 次元離散トーラス $(\mathbb{Z}/(N\mathbb{Z}))^q$ 上でのマルコフ過程モデルと線形方程式との大数の法則を用いた比較がなされた。[3],[4]は非線形熱方程式 $u' = \Delta u + R(u)$ に対して1次元離散トーラス $\mathbb{Z}/(N\mathbb{Z})$ 上での隣同士への移動のみのランダムウォーク(Kotelenezによって導入された確率モデルの $q=1$ の場合)による確率モデルが L^∞ の意味で確率収束することを示した。今回は隣同士への移動のみであったランダムウォークを離散トーラス上の全地点への移動を許すランダムウォークに拡張し、このデータ拡散の確率モデルが対応する非線形反応拡散方程式に L^∞ の意味で確率収束することを示す。この拡張によりネットワーク上でのデータ拡散を微分方程式で表現することが可能となる。

2. 非線形連続反応拡散モデル

任意の $x \in \mathbb{R}$ に対して生成・死滅を表す $b(x), d(x)$ をそれぞれ係数が非負値かつ $d(0) = 0$ なる m 次以下の多項式とし、非線形項を $R(x) = b(x) - d(x) = \sum_{i=0}^m c_i x^i$, $c_m < 0$ とする次のトーラス $[0, 1]$ 上の非線形反応拡散方程式を考える：

$$\begin{cases} \frac{d}{dt}u(t, r) = \int_0^1 u(t, r')dr' - u(t, r) + R(u(t, r)), \\ u(t, 0) = u(t, 1), \\ 0 \leq u(0, r) < \rho < \infty. \end{cases} \quad (2.1)$$

第一式右辺における第一項は周りから位置 r への拡散、第二項は位置 r から周りへの拡散、第三項は位置 r での生成・死滅を表している。さらに第二式、第三式はそれぞれ周期1の周期的境界条件であること、初期条件が発散していないことを意味している。 x が十分大きい場合($x > \rho$)は $R(x) < 0$ とする。(2.1)は積分方程式により完全グラフ上でのデータの生成・伝搬の流体モデルとして考えることができる。(2.1)の方程式の導出は[6]によってなされており、[2]により解の存在や一意性は報告されている。

$N \in \mathbb{N}$ に対してトーラス上の単関数族 $H^N = \left\{ f : [0, 1] \rightarrow \mathbb{R} \mid f(r) = \sum_{k=0}^{N-1} a_k \mathbb{1}_{[k/N, (k+1)/N)}(r) \right\}$ とおき、 $f \in H^N$ に対して

$$I_N f(r) = \frac{1}{N} \sum_{i=0}^{N-1} f(r + iN^{-1}) - f(r)$$

と定義し、(2.1)を空間離散化した方程式を考える：

$$\begin{cases} \frac{d}{dt}u^N(t, r) = I_N u^N(t, r) + R(u^N(t, r)), \\ u^N(t, 0) = u^N(t, 1), \\ u^N(0, r) = N \int_{kN^{-1}}^{(k+1)N^{-1}} u(0, r')dr' \text{ for } r \in [kN^{-1}, (k+1)N^{-1}). \end{cases} \quad (2.2)$$

ここで (2.1), (2.2) の解 u と u^N に対する次の結果が得られ、 u^N は u をうまく近似できることがわかる。

定理 2.1. u, u^N をそれぞれ (2.1), (2.2) の解とすると、

$$\sup_{[0, T]} \|u^N(t) - u(t)\|_\infty \leq C(T, R)N^{-1} \quad \text{for } T > 0$$

が成り立つ。

Proof. [5] の Proposition 4.1 参照. □

$f, g \in H^N$ に対して $L^2([0, 1])$ 内積は $\langle f, g \rangle = \sum_{k=0}^{N-1} f(kN^{-1})g(kN^{-1})N^{-1}$ で与える。
 $T_N(t) = \exp(tI_N)$ は I_N を生成作用素とする H^N 上の半群であり、任意の $f \in H^N$ に対して

$$T_N(t)f(r) = \frac{1}{N} \sum_{i=0}^{N-1} f(iN^{-1})(1 - e^{-t}) + f(r)e^{-t} \quad (2.3)$$

が成り立つ。また $T_N(t)I_N f = I_N T_N(t)f$ が成り立ち、 T_N, I_N は $(H^N, \langle \cdot, \cdot \rangle)$ 上の自己共役作用素であることがわかる。

3. 空間上の生成死滅拡散マルコフモデル

空間が離散化された (2.2) をさらに粒子化することを考える。 $N \geq 1$ を固定し、 $i \in \{0, \dots, k-1, k+1, \dots, N-1\}$ とする。時刻 t でのマルコフ連鎖の状態を $n(t) = (n_0(t), \dots, n_{N-1}(t)) \in \{\mathbb{N} \cup \{0\}\}^N$ とする。ここで $n_k(t)$ は時刻 t に場所 k に存在するデータ数を表している。次にデータの遷移レートを以下のように定義する：

$$\begin{cases} n \rightarrow n_{(i+1, k:-1)} = (\dots, n_i + 1, \dots, n_k - 1, \dots) & \text{at rate } n_k N^{-1}, \\ n \rightarrow n_{(k:+1)} = (\dots, n_k + 1, \dots) & \text{at rate } lb(n_k l^{-1}), \\ n \rightarrow n_{(k:-1)} = (\dots, n_k - 1, \dots) & \text{at rate } ld(n_k l^{-1}). \end{cases}$$

$n(t)$ はある確率空間上で定義された右連続左極限な関数とし、 $F_t^{N, l}$ を完備化された σ -加法族 $\sigma(n(s); s \leq t)$ 、 $\delta n(t) = n(t) - n(t-)$ とする。さらに $F_t^{N, l}$ -停止時刻 τ (つまり M への到達時刻) が存在して

$$\sup_{t \in [0, T]} \sup_k 1_{\{\tau > 0\}} n_k(t \wedge \tau) \leq M(T, N, l) < \infty$$

を満たすすると、次が成り立つ：

補題 3.1. 以下は mean 0 martingale である：

$$\begin{aligned} (a) \quad & n_k(t) - n_k(0) - \int_0^t \frac{1}{N} \sum_{i=0}^{N-1} (n_i(s) - n_k(s)) ds - \int_0^t lR(n_k(s)l^{-1}) ds. \\ (b) \quad & \sum_{s \leq t} (\delta n_k(s))^2 - \int_0^t \frac{1}{N} \sum_{i=0, i \neq k}^{N-1} (n_i(s) + n_k(s)) ds - \int_0^t \{lb(n_k(s)l^{-1}) + ld(n_k(s)l^{-1})\} ds. \\ (c) \quad & \sum_{s \leq t} (\delta n_k(s))(\delta n_i(s)) + \int_0^t \frac{1}{N} (n_i(s) + n_k(s)) ds \quad \text{for } \forall i \neq k. \end{aligned}$$

ここで (a) は場所 k での時刻 t までのデータ数の変化量、(b) は場所 k での事象数、(c) は場所 i, k 間の同時遷移事象数を表している。

以下のように Stochastic analogue を定義する：

$$X^N(t, r) = n_k(t)l^{-1} \quad \text{for } r \in [kN^{-1}, (k+1)N^{-1}).$$

補題 3.1(a) より

$$X^N(t) = X^N(0) + \int_0^t I_N X^N(s) ds + \int_0^t R(X^N(s)) ds + Z^N(t) \quad (3.1)$$

となる確率的誤差項を表す martingale $Z^N(t)$ が存在する.

4. 大数の法則によるモデル比較

次の定理が本研究の主結果である:

定理 4.1. (i) $\|X^N(0) - u(0)\|_\infty \rightarrow 0$ in probability as $N \rightarrow \infty$.

(ii) $\lim_{N \rightarrow \infty} \log(N/l) = 0$ を満たす $l = l(N)$ が存在する.

が成立するならば、任意の $T \geq 0$ に対して次が成り立つ:

$$\sup_{[0, T]} \|X^N(t) - u(t)\|_\infty \rightarrow 0 \text{ in probability as } N \rightarrow \infty.$$

証明のためにいくつか準備をする.

補題 4.2. $T_N(t)$ は $(H^N, \|\cdot\|_\infty)$ 上の正縮小半群である.

Proof. 任意の $f \in H^N$ に対して $\langle T_N(t)f, f \rangle \geq 0$ は明らか. さらに (2.3) に対して L^∞ ノルムをとると

$$\begin{aligned} \|T_N(t)f(r)\|_\infty &\leq \frac{1}{N} \left\| \sum_{i=0}^{N-1} f(iN^{-1}) \right\|_\infty (1 - e^{-t}) + \|f(r)\|_\infty e^{-t} \\ &\leq \frac{1}{N} \sum_{i=0}^{N-1} \|f(r)\|_\infty (1 - e^{-t}) + \|f(r)\|_\infty e^{-t} = \|f(r)\|_\infty. \end{aligned}$$

□

補題 4.3. $f = N \mathbb{1}_{[kN^{-1}, (k+1)N^{-1})}$ とおくと次が成り立つ:

$$\langle (T_N(t)f)^2, 1 \rangle - \langle T_N(t)f I_N T_N(t)f, 1 \rangle \leq h_N(t).$$

ここで $h_N(t)$ は $\int_0^t h_N(s) ds \leq t + N$ を満たす関数である.

Proof. $T_N(t), I_N$ の性質を用いると

$$\begin{aligned} \langle (T_N(t)f)^2, 1 \rangle &= \langle T_N(t)f, T_N(t)f \rangle = \langle T_N(2t)f, f \rangle = \langle (1 - e^{-2t}) + f e^{-2t}, f \rangle \\ &= (1 - e^{-2t}) + N e^{-2t} \\ &\leq N e^{-2t} + 1. \end{aligned}$$

$$\begin{aligned} -\langle T_N(t)f I_N T_N(t)f, 1 \rangle &= -\langle I_N T_N(t)f, T_N(t)f \rangle = -\langle T_N(t) I_N f, T_N(t)f \rangle \\ &= -\langle T_N(2t) I_N f, f \rangle \\ &= -\langle e^{-2t} - f e^{-2t}, f \rangle \\ &= (N - 1) e^{-2t} \\ &\leq N e^{-2t}. \end{aligned}$$

ゆえに $h_N(t) = 2N e^{-2t} + 1$ とおけば題意が満たされる.

□

次の補題は定理 4.1 の証明での誤差評価において重要である.

補題 4.4. $m(t)$ は $m(t_0) = 0$ である $[t_0, t_1]$ 上で定義された有界変動 martingale であり、さらに以下を満たす任意の確率過程とする:

(i) m は右連続左極限関数.

(ii) $t_0 \leq t \leq t_1$ に対して $|\delta m(t)| \leq 1$.

(iii) $0 \leq g(s) \leq h(s)$ に関して $\sum_{t_0 \leq s \leq t_1} (\delta m(s))^2 - \int_{t_0}^{t_1} g(s) ds$ は mean 0 martingale,

ここで $h(s)$ は決定関数、 $g(s)$ は $F_t^{N,l}$ - 適合 とする.

ならば、 $E[\exp(m(t_1))] \leq \exp(\frac{9}{2} \int_{t_0}^{t_1} h(s) ds)$.

Proof. [4] 参照. □

定理 4.1 の証明に入る. まず任意の $t \geq 0$ に対して $0 \leq u(t) < \rho$ 、 $\|X^N(0) - u(0)\|_\infty \rightarrow 0$ in probability という初期条件に対する仮定より

$$0 \leq X^N(0) < \rho \text{ for all } N$$

としても一般性は失われないことに注意する. また定理 2.1 より $\sup_{[0,T]} \|X^N(t) - u^N(t)\|_\infty$ について考えれば良い. 任意の $\epsilon_0 \in (0, 1]$ を固定し、 $\tau = \inf \{t : \|X^N(t) - u^N(t)\|_\infty > \epsilon_0\}$ とおき、

$$\bar{X}^N(t) = X^N(t \wedge \tau) + \int_{t \wedge \tau}^t I_N \bar{X}^N(s) ds + \int_{t \wedge \tau}^t R(\bar{X}^N(s)) ds \quad (4.1)$$

と定義する. (4.1) は $\bar{X}$ が時刻 $t \wedge \tau$ まで X の値をとり、その後は連続モデル的に変動することを意味している. τ の定義から

$$\begin{aligned} P(\sup_{[0,T]} \|X^N(t) - u^N(t)\|_\infty > \epsilon_0) &\leq P(\sup_{[0,T]} \|X^N(t \wedge \tau) - u^N(t \wedge \tau)\|_\infty \geq \epsilon_0) \\ &\leq P(\sup_{[0,T]} \|\bar{X}^N(t) - u^N(t)\|_\infty \geq \epsilon_0). \end{aligned}$$

ゆえに $\sup_{[0,T]} \|\bar{X}^N(t) - u^N(t)\|_\infty$ について考えれば良い. (3.1), (4.1) より

$$\bar{X}^N(t) = X^N(0) + \int_0^t I_N \bar{X}^N(s) ds + \int_0^t R(\bar{X}^N(s)) ds + Z^N(t \wedge \tau). \quad (4.2)$$

X^N , $\bar{X}^N$ の定義から

$$\|\delta \bar{X}^N(t)\|_\infty = \|\delta X^N(t \wedge \tau)\|_\infty \leq N^{-1}.$$

定理 2.1 より任意の $t \geq 0$ に対して $0 \leq u^N(t) < \rho$ であることを用いると

$$0 \leq X^N(t \wedge \tau) < \rho + 1 \text{ for } t \geq 0. \quad (4.3)$$

$\tau < t$ として (4.1) の両辺を t について微分することで次の微分方程式が得られる:

$$\frac{d}{dt} \bar{X}^N(t) = I_N \bar{X}^N(t) + R(\bar{X}^N(t)).$$

この方程式を解くと

$$\bar{X}^N(t) = T_N(t - \tau) X^N(\tau) + \int_\tau^t T_N(t - s) R(\bar{X}^N(s)) ds.$$

ゆえに補題 4.2 と R の定義から背理法を用いることで

$$0 \leq \bar{X}^N(t) < \rho + 1 \text{ for } t \geq 0 \quad (4.4)$$

が成り立つ. さらに (4.2) から得られる微分方程式と (2.2) より

$$\frac{d}{dt}(\bar{X}^N(t) - u^N(t)) = I_N(\bar{X}^N(t) - u^N(t)) + R(\bar{X}^N(t)) - R(u^N(t)) + \frac{d}{dt}Z^N(t \wedge \tau)$$

が得られる. この方程式を解くと

$$\bar{X}^N(t) - u^N(t) = T_N(t)(X^N(0) - u^N(0)) + \int_0^t T_N(t-s)(R(\bar{X}^N(t)) - R(u^N(t)))ds + \int_0^t T_N(t-s)dZ^N(s \wedge \tau).$$

今後、 $Y^N(t) = \int_0^t T_N(t-s)dZ^N(s \wedge \tau)$ とかく. $Y^N(t)$ は確率的誤差の線形的な時間発展を表している. ここで両辺の L^∞ ノルムをとると

$$\begin{aligned} \|\bar{X}^N(t) - u^N(t)\|_\infty &\leq \|T_N(t)(X^N(0) - u^N(0))\|_\infty + \left\| \int_0^t T_N(t-s)(R(\bar{X}^N(t)) - R(u^N(t)))ds \right\|_\infty + \|Y^N(t)\|_\infty \\ &\leq \|X^N(0) - u^N(0)\|_\infty + \int_0^t \|R(\bar{X}^N(t)) - R(u^N(t))\|_\infty ds + \|Y^N(t)\|_\infty \\ &\leq \|X^N(0) - u^N(0)\|_\infty + K \int_0^t \|\bar{X}^N(t) - u^N(t)\|_\infty ds + \|Y^N(t)\|_\infty. \end{aligned}$$

2つ目の不等式では補題 4.2、3つ目の不等式では R が m 次の多項式であることと (4.4) を用いた. ここで K は ρ と R の係数に依存するような定数である. よって Gronwall の不等式を適用すると

$$\|\bar{X}^N(t) - u^N(t)\|_\infty \leq \{ \|X^N(0) - u^N(0)\|_\infty + \|Y^N(t)\|_\infty \} \exp(Kt).$$

よって次の補題を示せば定理 4.1 が証明されたことになる.

補題 4.5. $\sup_{[0,T]} \|Y^N(t)\|_\infty \rightarrow 0$ in probability as $N \rightarrow \infty$.

Proof. $\bar{t} \in (0, T]$, $k \in \{0, 1, \dots, N-1\}$ を固定し、 $f = N\mathbb{1}_{[kN^{-1}, (k+1)N^{-1})}$ とおく. さらに $0 \leq t \leq \bar{t}$ に対して $\bar{m}(t) = \langle \int_0^t T_N(\bar{t}-s)dZ^N(s \wedge \tau), f \rangle$ とする. ここで $\bar{m}$ は $0 \leq t \leq \bar{t}$ 上で mean 0 martingale、 $\bar{m}(\bar{t}) = Y^N(\bar{t}, kN^{-1})$ であることに注意する. $\delta Z^N(t, kN^{-1}) = l^{-1}\delta n_k(t)$ であることを用いると、任意の $\phi \in H^N$ に対して

$$\begin{aligned} &\sum_{s \leq t} (\delta \langle Z^N(s \wedge \tau), \phi \rangle)^2 \\ &= \sum_{s \leq t} \left(\frac{1}{N} \sum_{k=0}^{N-1} \delta Z^N(s \wedge \tau, kN^{-1}) \phi(kN^{-1}) \right)^2 \\ &= (Nl)^{-2} \sum_{s \leq t} \left(\sum_{k=0}^{N-1} \delta n_k(s \wedge \tau) \phi(kN^{-1}) \right)^2 \\ &= (Nl)^{-2} \sum_{s \leq t} \left\{ \sum_{k=0}^{N-1} (\delta n_k(s \wedge \tau))^2 \phi(kN^{-1})^2 + \sum_{\substack{i, k=0 \\ i \neq k}}^{N-1} \delta n_k(s \wedge \tau) \delta n_i(s \wedge \tau) \phi(kN^{-1}) \phi(iN^{-1}) \right\}. \end{aligned}$$

上式に対して補題 3.1(b),(c) を用いることで、

$$\begin{aligned} &\sum_{s \leq t} (\delta \langle Z^N(s \wedge \tau), \phi \rangle)^2 - (Nl)^{-1} \int_0^{t \wedge \tau} \{ -\langle X^N(s), \phi I_N \phi \rangle + \langle \phi^2, I_N X^N(s) \rangle \\ &\quad - \langle \phi, I_N(X^N(s)\phi) \rangle + \langle b(X^N(s)) + d(X^N(s)), \phi^2 \rangle \} ds \end{aligned}$$

は mean 0 martingale であることがわかる. ゆえに $\phi = T_N(\bar{t} - s)f$ とおけば $0 \leq t \leq \bar{t}$ に対して

$$\begin{aligned} \sum_{s \leq t} (\delta \bar{m}(s))^2 - (Nl)^{-1} \int_0^{t \wedge \tau} \{ -\langle X^N(s), T_N(\bar{t} - s)f I_N T_N(\bar{t} - s)f \rangle + \langle (T_N(\bar{t} - s)f)^2, I_N X^N(s) \rangle \\ - \langle T_N(\bar{t} - s)f, I_N(X^N(s)T_N(\bar{t} - s)f) \rangle + \langle b(X^N(s)) + d(X^N(s)), (T_N(\bar{t} - s)f)^2 \rangle \} ds \end{aligned} \quad (4.5)$$

も mean 0 martingale であることがわかる. $\theta \in [0, 1]$ に対して $m(t) = \theta l \bar{m}(t)$ とおくと $|\delta \bar{m}(s)| \leq l^{-1}$ であることから $|\delta m(t)| = \theta l |\delta \bar{m}(s)| \leq \theta \leq 1$ が成り立つ. $m(t)$ の取り方から (4.5) より

$$\begin{aligned} \sum_{s \leq t} (\delta m(s))^2 - (Nl)^{-1} (\theta l)^2 \int_0^{t \wedge \tau} \{ -\langle X^N(s), T_N(\bar{t} - s)f I_N T_N(\bar{t} - s)f \rangle + \langle (T_N(\bar{t} - s)f)^2, I_N X^N(s) \rangle \\ - \langle T_N(\bar{t} - s)f, I_N(X^N(s)T_N(\bar{t} - s)f) \rangle + \langle b(X^N(s)) + d(X^N(s)), (T_N(\bar{t} - s)f)^2 \rangle \} ds \end{aligned}$$

も mean 0 martingale である. ここで

$$\begin{aligned} g(s) = (Nl)^{-1} (\theta l)^2 \{ -\langle X^N(s), T_N(\bar{t} - s)f I_N T_N(\bar{t} - s)f \rangle + \langle (T_N(\bar{t} - s)f)^2, I_N X^N(s) \rangle \\ - \langle T_N(\bar{t} - s)f, I_N(X^N(s)T_N(\bar{t} - s)f) \rangle + \langle b(X^N(s)) + d(X^N(s)), (T_N(\bar{t} - s)f)^2 \rangle \} \end{aligned}$$

とおくと, (4.3) と補題 4.3 より ρ を含む定数項を $c(\rho)$ とすると $g(s) \leq c(\rho)(Nl)^{-1} (\theta l)^2 h_N(t)$ なる関数 $h_N(t)$ が存在する. 従って補題 4.4 を適用することで

$$\mathbb{E}[\exp(\theta l \bar{m}(\bar{t}))] \leq \exp[c(\rho) \theta^2 l (1 + \bar{t} N^{-1})].$$

ここで $\bar{t} \leq T$ から $\bar{t}/N \leq 1$. ゆえに任意の $\epsilon > 0$ に対して

$$\begin{aligned} P(Y^N(\bar{t}, kN^{-1}) > \epsilon) &= P(\theta l Y^N(\bar{t}, kN^{-1}) > \theta l \epsilon) \\ &\leq \mathbb{E}[\exp(\theta l Y^N(\bar{t}, kN^{-1}))] \exp(-\theta l \epsilon) \\ &\leq \exp(\theta l (c(\rho) \theta - \epsilon)). \end{aligned}$$

従って $N, l, k, \bar{t}$ に依存しない θ が以下を満たすようにとれる:

$$P(Y^N(\bar{t}, kN^{-1}) > \epsilon) \leq \exp(-a \epsilon^2 l) \quad \text{for } a = a(\rho) > 0.$$

$-Y^N(\bar{t}, kN^{-1})$ に対しても同様に考えることで

$$P(-Y^N(\bar{t}, kN^{-1}) > \epsilon) \leq \exp(-a \epsilon^2 l).$$

これらの結果をまとめると任意の $t \in (0, T]$, $k \in \{0, 1, \dots, N-1\}$ に対して

$$P(|Y^N(\bar{t}, kN^{-1})| > \epsilon) \leq 2 \exp(-a \epsilon^2 l)$$

が成り立つ. ここで $\|Y^N(t)\|_\infty = \sup_k |Y^N(t, kN^{-1})|$, $Y(0) = 0$ より任意の $0 \leq t \leq T$ に対して

$$\begin{aligned} P(\|Y^N(\bar{t}, kN^{-1})\|_\infty > \epsilon) &= P\left(\sup_k |Y^N(\bar{t}, kN^{-1})| > \epsilon\right) \\ &\leq \sum_k P(|Y^N(\bar{t}, kN^{-1})| > \epsilon) \leq 2N \exp(-a \epsilon^2 l) \end{aligned} \quad (4.6)$$

が成り立つ.

次に $\|Y^N(t)\|_\infty$ を $\sup_{[0, T]} \|Y^N(t)\|_\infty$ に取り替えた場合を考える. $Y^N(t) = \int_0^t I_N Y^N(s) ds + Z^N(t \wedge \tau)$ より任意の $nTN^{-1} \leq t \leq (n+1)TN^{-1}$, $0 \leq n \leq N-1$ に対して $\tilde{m}(t) = Z^N(t \wedge \tau) - Z^N(nTN^{-1} \wedge \tau)$

とおくと

$$Y^N(t) = Y^N(nTN^{-1}) + \int_{nTN^{-1}}^t I_N Y^N(s) ds + \tilde{m}(t).$$

上式の両辺に対して L^∞ ノルムをとると

$$\begin{aligned} \|Y^N(t)\|_\infty &\leq \|Y^N(nTN^{-1})\|_\infty + \int_{nTN^{-1}}^t \|I_N Y^N(s)\|_\infty ds + \|\tilde{m}(t)\|_\infty \\ &\leq \|Y^N(nTN^{-1})\|_\infty + 2N \int_{nTN^{-1}}^t \|Y^N(s)\|_\infty ds + \|\tilde{m}(t)\|_\infty. \end{aligned}$$

ここで Gronwall の不等式を適用することで

$$\|Y^N(t)\|_\infty \leq \{\|Y^N(nTN^{-1})\|_\infty + \|\tilde{m}(t)\|_\infty\} e^{2N(t-nTN^{-1})}$$

が得られる. ゆえに両辺において $[nTN^{-1}, (n+1)TN^{-1}]$ 上の上限をとると

$$\sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|Y^N(t)\|_\infty \leq \left\{ \|Y^N(nTN^{-1})\|_\infty + \sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|\tilde{m}(t)\|_\infty \right\} e^{2T}. \quad (4.7)$$

$k \in \{0, 1, \dots, N-1\}$, $\theta \in [0, 1]$ を固定し, $m(t) = \theta l \tilde{m}(t, kN^{-1})$ とおくと, 補題 3.1(b) と $\delta Z^N(t, kN^{-1}) = \delta X^N(t, kN^{-1}) = l^{-1} \delta n_k(t)$ より任意の $t \in [nTN^{-1}, (n+1)TN^{-1}]$ に対して

$$\begin{aligned} \sum_{nTN^{-1} \leq s \leq t} (\delta m(s))^2 - \theta^2 l \int_{nTN^{-1} \wedge \tau}^{t \wedge \tau} \frac{1}{N} \sum_{\substack{i, k=0 \\ i \neq k}} (X^N(s, iN^{-1}) + X^N(s, kN^{-1})) ds \\ - \theta^2 l \int_{nTN^{-1} \wedge \tau}^{t \wedge \tau} \{b(X^N(s, kN^{-1})) + d(X^N(s, kN^{-1}))\} ds \end{aligned}$$

は mean 0 martingale である. ここで

$$g(s) = \theta^2 l \frac{1}{N} \sum_{\substack{i, k=0 \\ i \neq k}} (X^N(s, iN^{-1}) + X^N(s, kN^{-1})) + \theta^2 l \{b(X^N(s, kN^{-1})) + d(X^N(s, kN^{-1}))\}$$

とおくと, (4.3) よりある定数 $c(\rho)$ が存在して $g(s) \leq c(\rho) \theta^2 l N$ が成り立つ. ゆえに $|\delta m(t)| \leq 1$ であることから補題 4.4 を適用すると

$$\mathbb{E}[\exp(m((n+1)TN^{-1}))] \leq \exp\left(\frac{9}{2} \int_{nTN^{-1}}^{(n+1)TN^{-1}} c(\rho) \theta^2 l N ds\right) = \exp(c(\rho) \theta^2 l T).$$

ここで $\tilde{m}$ は martingale より Doob の不等式を適用すれば

$$\begin{aligned} P\left(\sup_{[nTN^{-1}, (n+1)TN^{-1}]} \tilde{m}(t, kN^{-1}) \geq 0\right) &\leq E[\exp(m((n+1)TN^{-1}))] \exp(-\theta l \epsilon) \\ &\leq \exp(\theta l (c(\rho) T \theta - \epsilon)). \end{aligned}$$

ゆえに以下を満たすような $N, l, k, \bar{t}$ に依存しない θ が存在する:

$$P\left(\sup_{[nTN^{-1}, (n+1)TN^{-1}]} \tilde{m}(t, kN^{-1}) \geq 0\right) \leq \exp(-a\epsilon^2 l) \quad \text{for } a = a(\rho, T) > 0$$

$-\tilde{m}(t, kN^{-1})$ に対しても同様に計算することで

$$P \left(\sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|\tilde{m}(t)\|_{\infty} \right) \leq 2N \exp(-a\epsilon^2 l) \quad (4.8)$$

が得られる. 従って (4.6), (4.7), (4.8) より

$$\begin{aligned} P \left(e^{-4T} \sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|Y^N(t)\|_{\infty} > \epsilon \right) &\leq P \left(\|Y^N(\bar{t}, kN^{-1})\|_{\infty} > \epsilon \right) + P \left(\sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|\tilde{m}(t)\|_{\infty} \right) \\ &\leq 4N \exp(-a\epsilon^2 l) \end{aligned}$$

が成り立つ. また

$$\begin{aligned} P \left(e^{-4T} \sup_{[0, T]} \|Y^N(t)\|_{\infty} > \epsilon \right) &= P \left(e^{-4T} \bigcup_n \sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|Y^N(t)\|_{\infty} > \epsilon \right) \\ &\leq \sum_{n=0}^{N-1} P \left(e^{-4T} \sup_{[nTN^{-1}, (n+1)TN^{-1}]} \|Y^N(t)\|_{\infty} > \epsilon \right) \\ &\leq 4N^2 \exp(-a\epsilon^2 l). \end{aligned}$$

これは定理 4.1 の条件 (ii) より, $\sup_{[0, T]} \|Y^N(t)\|_{\infty}$ が 0 に確率収束していることを意味している. $\square$

参考文献

- [1] L.Arnold, M.Theodosopulu, *deterministic limit of the stochastic model of chemical reactions with diffusion*, Adv.in Appl.probab.12, (1980)367-379
- [2] P.W.Bates, G.Zhao, *Existence, uniqueness and stability of the stationary solutions to a nonlocal evolution equation arising in population dispersal*, J.Math.Anal.Appl.332, (2007)428-440
- [3] D.Blount, *Comparison of stochastic model of a chemical reaction with diffusion and the deterministic model*, Ph.D.dissertation, Univ. Wisconsin-Madison, (1987)
- [4] D.Blount, *Law of large numbers in the supremum norm for a chemical reaction with diffusion*, The Annals of Applied Probability, Vol.2, No.1, (1992)131-141
- [5] T.Franco, P.Groisman, *A particle system with explosion: Law of large numbers for the density of particles and the blow-up time*, Journal of Statistical Physics, Volume149, Issue4, (2012)629-642
- [6] V.Hutson, S.Martinez, K.Mischaikow, G.T.Vickers, *The evolution of dispersal*, J.Math.Biol, 47, (2003)487-517
- [7] P.Kotelenez, Ph.D.dissertation, Universität Bremen Forschungsschwerpunkt Dynamische Systemes, (1982)
- [8] P.Kotelenez, *A submartingale type inequality with applications to stochastic evolution equations*, Stochastic 8, (1982)139-151
- [9] P.Kotelenez, *Law of large numbers and central limit theorem for linear chemical reactions with diffusion*, Ann.Probab, 14, (1986)352-370
- [10] P.Kotelenez, *Gaussian approximation to the nonlinear reaction-diffusion equation*, Report 146, Universität Bremen Forschungsschwerpunkt Dynamische Systemes, (1986)
- [11] P.Kotelenez, *Fluctuations near homogeneous states of chemical reactions with diffusion*, Adv.in Appl.Probab, 19, (1987)352-370
- [12] P.Kotelenez, *High density limit theorem for nonlinear chemical reactions with diffusion*, Probab.Theory Related Fields, 78, (1988)11-37

再試行型待ち行列を用いた 単一チャネルコグニティブ無線の解析

阿久津 康平
筑波大学大学院
システム情報工学研究科

Phung-Duc Tuan
筑波大学
システム情報系

1 はじめに

近年、スマートフォンやタブレット型端末等の情報通信端末の爆発的な増加が続いている。その結果、無線資源の周波数、とりわけ主に通信で用いられる周波数の不足が問題となる状況が起こっている [1]。そこで必要となってくるのが、周波数の利用効率を高める方法を生み出すことである。この方法の1つとして、考案されている技術がコグニティブ無線通信技術である。コグニティブ無線通信技術を利用することで、無線機器は各周波数帯において、他のシステムに干渉を与えない限り、他のシステムと周波数を共用して利用することが可能になる [2]。この技術を実現するにあたり、「周波数共有型」・「ヘテロジーニアス型」という2つの方法が検討されている [3]。「周波数共有型」は、あるサービス中の既存のシステムにおいて、空いている周波数帯、もしくは利用されていない時間帯が存在する周波数帯を無線機器が探し、必要な周波数帯を確保するという方法である。この方法が有効な例としては、ホワイトスペースが挙げられる [4]。ホワイトスペースとは、放送用など、ある目的のために周波数が割り当てられているが、地理的条件や時間的条件等の制約により使用されていない周波数帯のことである。このホワイトスペースを有効活用することにより、圧迫されつつある通信の周波数帯の余裕を持たせることができる。一方で、「ヘテロジーニアス型」は、無線機器が通信の混雑度や接続性を認識し、最適な通信方法を自動的に変更しながら通信を継続させるという方法である。この方法はスマートフォンなどの通信機器における通信網や無線 LAN トラフィックの切り替えなどに利用されている。さらには今後、第5世代携帯電話システムにおいても活用が期待されている [5]。両者の方法で用いられている、無線機器が周囲の状況を検知することをセンシングという。

一般に、コグニティブ無線では、優先権のある一次利用者 (以下、PU とする) と、優先権を持たない二次利用者 (以下、SU とする) といった複数種類のクラスのユーザが存在する。これらのユーザの権利の差を、待ち行列モデルを用いてモデル化した先行研究としては、次のような文献がある。Konishi [6] では、SU に割り当てるチャネル数が可変であるモデルに対し、SU のブロッキング確率や強制退去確率、スループット等の性能解析を行っている。また、Salameh [7] では、反応的決定

スペクトラムハンドオフを用いたコグニティブ無線における SU の性能評価を行い、センシングで
きる SU の人数の上限を設けた上で、SU がチャンネルにたどり着くまでの所要時間や、SU の遮断回
数及び遮断確率といった性能指標を評価している。[6] では、利用されていないチャンネル数が、SU
が要求したチャンネル数に満たないとき、SU はシステムから退去するモデルである。また、[7] では、
SU が 1 つずつチャンネルにセンシングを行い、全てのチャンネルが利用されているときにも、SU は少
なくとも 1 つのチャンネルが空くまで、同様のセンシングを行い続けるというモデルである。これら
のモデルに対し、本稿のモデルは、SU がシステムに初めて入ったとき、もしくはチャンネルに入ろう
としてチャンネルに空きがないとき、センシングを行うが、その際には SU はセンシングを行うため
のある仮想的な場所 (Sensing pool) に入り、センシングを行うモデルになっている。

本稿では、「周波数共有型」のコグニティブ無線のモデル化を行った。使用したモデルは、単一
サーバの再試行型待ち行列モデルである。このモデルを用いた理由は、SU のセンシング時間を考
慮するためである。再試行型待ち行列モデルでは、サーバに入ることができなかったユーザが、再
びサーバに入るために待機する orbit と呼ばれる場所を設定する。先に述べた仮想的な場所が本稿
のモデルにおいては orbit に該当する。このモデルを用いた先行研究としては、Keilson [8] が挙げ
られる。[8] では、サーバが稼働中の場合と非稼働中の場合の orbit 内の系内人数の同時確率分布
やサーバの状態が既知のときの系内人数の確率分布および各場合の母関数を導出した。また Falin
[9, 10] では、orbit 内のユーザの待ち時間に関する指標を導出し、Falin [11] では、高密度領域にお
ける orbit 内のユーザの待ち時間の極限定理を導いた。さらに、Falin [12] では、ユーザが再試行す
る回数に着目し、再試行回数の平均・分散・確率分布を導出している。今回列举した論文中のモデ
ルは、ユーザが単一クラスであるとしている。一方、本稿で扱うモデルでは、ユーザは PU と SU の
2 種類のクラスのユーザが存在するものとする。これらのユーザはそれぞれ異なる方法でサーバに
入り、サービスを受ける。

本稿は次のように構成される。第 2 節では、モデルの概要について述べる。第 3 節では、モデルを
解析する上で必要となる設定をおく。第 4 節で、orbit 内の SU の系内人数を評価する。第 5 節では、
第 4 節で導出した SU の系内人数の平均及び分散を、PU の到着率のみ変化させた場合、SU の到着
率のみ変化させた場合、PU と SU の到着率の両方を変化させた場合をグラフで表し、考察を行う。
最後に、第 6 節で、本稿で得られた結果及び今後の課題についてまとめる。

2 モデルの概要

今回のモデルでは、PU と SU の 2 種類のユーザを考える。PU と SU は、次の点で異なる挙動を
とる。(1) SU は、チャンネルをセンシングをするため、チャンネルに向かうまでに PU より時間を要す
る。一方で、PU はシステムに入ると直ちにチャンネルに向かう。(2) SU がチャンネルを占拠している
とき、PU がチャンネルに到達すると、SU を退去させ PU がチャンネルを占拠する。今回のモデルでは、
SU が orbit に滞在する時間を、センシングを行う時間とみなす。次にコグニティブ無線システム内
での PU と SU の挙動について述べる。PU の挙動は以下の手順で行われる。

1. システムに到着し、必要とする周波数を要求する。
2. 要求した周波数に、他のユーザが存在しない場合は、そのまま通信を開始する。要求した周波数に他のユーザが存在していて、通信中のユーザが PU の場合は、通信できないため、システムから退出する。一方、通信中のユーザが SU の場合は、SU を退出させ、通信を開始する。
3. 要求した周波数を用いて、通信を行う。
4. 通信終了後、システムから退出する。

また SU の挙動は以下の通りである。

1. システムに到着する。
2. 利用できる周波数を検知するために、センシングを始める。
3. センシング終了後、必要とする周波数を要求する。
4. 要求した周波数において、他のユーザが通信中の場合は、再びセンシングを行う。以降、周波数を確保できるまで、手順 2 からやり直す。
5. 要求した周波数を用いて、通信を行う。
6. 通信中に PU がチャネルにやってきた場合は一旦退去し、orbit に戻り、手順 2 からやり直す。
7. 通信終了後、システムから退出する。

上記の挙動を踏まえ、本稿では以下の仮定の下でシステムのモデル化を行う。PU の挙動は、

1. 率 λ_1 のポアソン過程に従って到着し、直接サーバに入る。
2. PU がサーバに入ろうとする際、サーバが稼動していない場合はそのままサーバに入る。サーバが稼動していて、PU がサービスを受けている場合はサービスを受けずに退出し、SU が

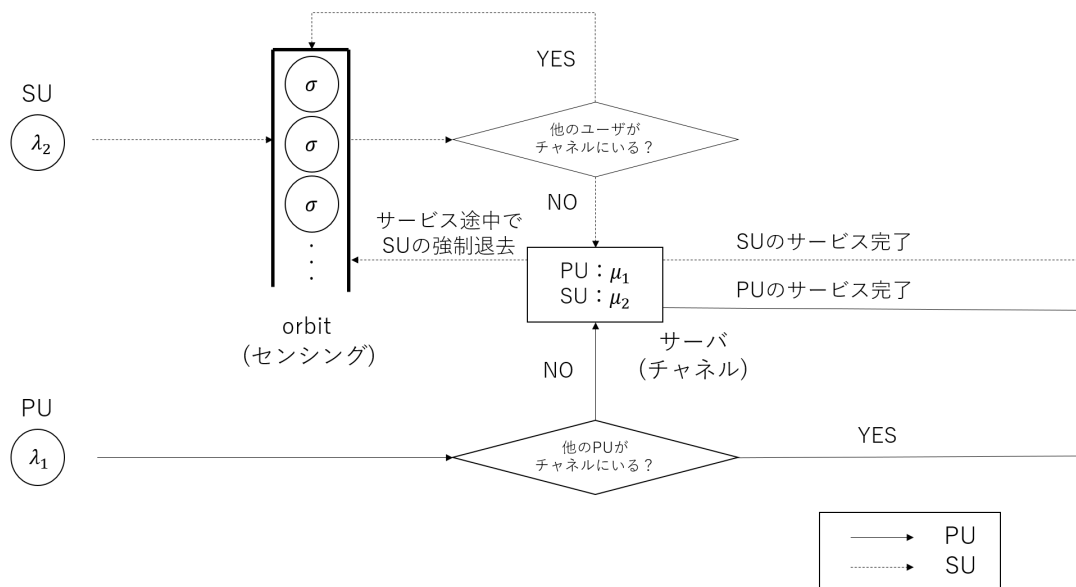


図 1 モデル図.

サービスを受けている場合は, その SU を orbit に移動させサーバに入る.

3. サーバに入るとサービス時間がパラメータ μ_1 の指数分布に従いサービスを受ける.
4. サービス終了後, 退出する.

また, SU の挙動は,

1. 率 λ_2 のポアソン過程に従って到着する.
2. orbit に入る.
3. orbit 内の各 SU は, 他の SU とは独立に, 率 σ のポアソン過程に従ってサーバに入る. つまり, orbit 内に j 人の SU が滞在しているときは, SU 全体では率 $j\sigma$ のポアソン過程に従ってサーバに入る.
4. サーバに入り, サーバが稼働している場合, orbit に戻る. 以降, サーバが非稼働の状態のときに入るまで, 手順 2 からやり直す.
5. サービス時間がパラメータ μ_2 の指数分布に従いサービスを受ける.
6. サービスを受けている最中に, PU がサーバにやってきた場合は orbit に戻り, 手順 2 からやり直す.
7. サービス終了後, 退出する.

モデルの注意点としては, 通信の途中で中断された SU の通信時間についてである. 途中までチャンネルに滞在し, 中断されたときまで行った通信時間は, 再びチャンネルで通信を行う際には考慮しない. これは, SU のサービス時間が指数分布に従うため, 自然な結果である. 図 1 は, 今回のモデルをフローチャートで表したものである.

3 モデルの設定

上記のモデルを解析するにあたり, 次の確率変数を定める. orbit 内に滞在している SU の人数を $N(t) \in \{0, 1, 2, \dots\}$, チャンネルの状態を $C(t) \in \{0, 1, 2\}$ とおく. チャンネルの状態が $C(t) = 0$ のときは, 誰も通信を行っていない状態を, $C(t) = 1$ は PU が通信を行っている状態を, $C(t) = 2$ は SU が通信を行っている状態を表している. このとき, $\{(C(t), N(t)) | t \geq 0\}$ は連続時間マルコフ連鎖となる. また, 定常状態及び PU と SU の入力過程, SU が再試行を行う時間間隔, PU と SU のサービス時間は互いに独立であると仮定する. さらに, チャンネルの状態が i ($i \in \{0, 1, 2\}$), orbit 内の SU の滞在人数が j ($j \in \mathbb{N} \cup \{0\}$) 人のときの結合確率を $\pi_{i,j} = P(C(t) = i, N(t) = j)$ とする. このとき, モデルが定常状態であると仮定しているため, 確率分布において時刻を考慮する必要がないことに注意する. また, $\Pi_i(z) := \sum_{j=0}^{\infty} \pi_{i,j} z^j$ とおく. 以上の設定の下で, モデルを解析していく.

4 評価指標の導出

ここでは、モデルの安定条件や orbit 内の SU の系内人数の平均及び分散、さらに各チャネルの状態における母関数を導出する。まず、平衡方程式を導出する。\$(C(t), N(t))\$ の推移に注意すると、以下の通りになる。ただし、\$\pi_{0,-1} = \pi_{1,-1} = \pi_{2,-1} = 0\$ と定める。

$$\begin{cases} (\lambda_1 + \lambda_2 + j\sigma)\pi_{0,j} = \lambda_2\pi_{0,j-1} + \mu_1\pi_{1,j} + \mu_2\pi_{2,j}, & (1) \\ (\lambda_2 + \mu_1)\pi_{1,j} = \lambda_1\pi_{0,j} + \lambda_1\pi_{2,j-1} + \lambda_2\pi_{1,j-1}, & (2) \\ (\lambda_1 + \lambda_2 + \mu_2)\pi_{2,j} = (j+1)\sigma\pi_{0,j+1} + \lambda_2\pi_{2,j-1}. & (3) \end{cases}$$

(1), (2) 及び (3) の各両辺に \$z^j\$ を掛けて \$j = 0, 1, 2, \dots\$ と和をとると、

$$\begin{cases} (\lambda_1 + \lambda_2)\Pi_0(z) + \sigma z\Pi_0'(z) = \lambda_2 z\Pi_0(z) + \mu_1\Pi_1(z) + \mu_2\Pi_2(z), & (4) \\ (\lambda_2 + \mu_1)\Pi_1(z) = \lambda_1\Pi_0(z) + \lambda_1 z\Pi_2(z) + \lambda_2 z\Pi_1(z), & (5) \\ (\lambda_1 + \lambda_2 + \mu_2)\Pi_2(z) = \sigma\Pi_0'(z) + \lambda_2 z\Pi_2(z), & (6) \end{cases}$$

となる。(5) 及び \$\Pi_0(1) + \Pi_1(1) + \Pi_2(1) = 1\$ より、

$$\Pi_1(1) = \frac{\lambda_1}{\mu_1 + \lambda_1}, \quad (7)$$

$$\Pi_0(1) + \Pi_2(1) = \frac{\mu_1}{\mu_1 + \lambda_1}. \quad (8)$$

また、(5) 及び (6) を変形して、

$$\Pi_1(z) = \frac{\lambda_1\Pi_2(z) + \lambda_1 z\Pi_2(z)}{\lambda_2 + \mu_1 - \lambda_2 z}, \quad (9)$$

$$\Pi_2(z) = \frac{\sigma\Pi_0'(z)}{\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z}. \quad (10)$$

(9) 及び (10) を (4) に代入すると、

$$\frac{\Pi_0'(z)}{\Pi_0(z)} = \gamma(z), \quad (11)$$

となる。ただし、

$$\gamma(z) := \frac{\lambda_2(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2 z)(\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z)}{\sigma(\mu_2(\lambda_2 + \mu_1) - \lambda_2(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)z + z^2\lambda_2^2)}, \quad (12)$$

である。すると、(10), (11) 及び (12) より、

$$\Pi_2(1) = \frac{\lambda_2(\lambda_1 + \mu_1)}{\mu_2\mu_1 - \lambda_1\lambda_2 - \lambda_2\mu_1}\Pi_0(1). \quad (13)$$

(8), (13) より、

$$\Pi_0(1) = \frac{\mu_2\mu_1 - \lambda_1\lambda_2 - \lambda_2\mu_1}{\mu_2(\mu_1 + \lambda_1)}, \quad (14)$$

$$\Pi_2(1) = \frac{\lambda_2}{\mu_2}. \quad (15)$$

このとき, (7), (14), (15) において, $\Pi_0(1) = P(C(t) = 0)$, $\Pi_1(1) = P(C(t) = 1)$, $\Pi_2(1) = P(C(t) = 2)$ であり, これらが確率分布となるために, $0 < P(C(t) = i) < 1$ ($i \in \{0, 1, 2\}$) とならなければならない. このことから, 安定条件の必要条件として以下を得ることができる.

補題 1. 安定条件の必要条件は

$$\frac{\lambda_2}{\mu_2} + \frac{\lambda_1}{\mu_1 + \lambda_1} < 1, \quad (16)$$

$$\Leftrightarrow \frac{\lambda_2}{\mu_2} < \frac{\mu_1}{\mu_1 + \lambda_1}, \quad (17)$$

で得られる.

補題の条件式の考察を行う. (16) からは, PU のサービス率に比べ, PU の到着率が十分に大きいときには左辺の第 2 項が 1 に近づくため, 条件を満たすためには第 1 項が小さくならなければならない, つまり SU の到着率が SU のサービス率よりも十分小さくなくてはならないということが分かる. また, (17) からは, PU の到着率に比べ, PU のサービス率が十分大きいときには右辺が 1 に近づくため, このときの SU は M/M/1 型待ち行列に近い安定条件を満たせばよいということが分かる. そして, (7) より, (17) の右辺が PU がチャネルを使用していない確率であることが分かるが, このことと左辺が SU のトラヒック密度であることに注意すると, チャネルが PU により使用されていない確率よりも SU のトラヒック密度は小さいということが分かる. また, この条件には σ が現れないことが分かる. このことから, チャネルの状態は再試行率に依存せず, PU と SU の到着率及びサービス率のみに依存することが分かる. 以下, 補題 1 の条件の下で進めていく. (11) を解くと, 以下の結果が得られる.

$$\Pi_0(z) = \Pi_0(1) \exp\left(\int_1^z \gamma(u) du\right). \quad (18)$$

$\gamma(z)$ の極を z_1, z_2 ($z_1 < z_2$) とおくと,

$$z_1 = \frac{\lambda_1 + \lambda_2 + \mu_1 - \lambda_2 - \sqrt{(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2)^2 - 4\mu_1(\lambda_2 + \mu_1)}}{2\lambda_2},$$

$$z_2 = \frac{\lambda_1 + \lambda_2 + \mu_1 - \lambda_2 + \sqrt{(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2)^2 - 4\mu_1(\lambda_2 + \mu_1)}}{2\lambda_2}.$$

ここで, (16) を満たすことと, $1 < z_1$ となることは同値であることに注意する. このとき, $\gamma(z)$ は

$$\gamma(z) = \frac{A}{z_1 - z} + \frac{B}{z_2 - z}. \quad (19)$$

ただし,

$$A = \frac{\lambda_2}{\sigma} \frac{(\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z_1)(\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z_1)}{\lambda_2^2(z_2 - z_1)},$$

$$B = \frac{\lambda_2}{\sigma} \frac{(\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z_2)(\lambda_1 + \lambda_2 + \mu_2 - \lambda_2 z_2)}{\lambda_2^2(z_1 - z_2)},$$

である。(9), (10) より,

$$\Pi_1'(1) = \frac{\mu_1(\lambda_1 \Pi_0'(1) + \lambda_2 \Pi_2(1) + \lambda_1 \Pi_2'(1)) + \lambda_2 \lambda_1 (\Pi_0(1) + \Pi_2(1))}{\mu_1^2}, \quad (20)$$

$$\Pi_2'(1) = \frac{\sigma \Pi_0''(1)(\lambda_1 + \mu_2) + \lambda_2 \sigma \Pi_0'(1)}{(\lambda_1 + \mu_2)^2}. \quad (21)$$

これらはそれぞれ, チャンネルに PU がいるときの orbit 内の SU の系内人数の平均と, チャンネルに SU がいるときの orbit 内の SU の系内人数の平均である. また (11) より,

$$\Pi_0'(1) = \frac{\lambda_2(\lambda_1 + \mu_2)}{\sigma \mu_2}, \quad (22)$$

$$\Pi_0''(1) = \gamma'(1) \Pi_0(1) + \gamma(1) \Pi_0'(1). \quad (23)$$

(20), (21), (22) より, $\Pi_0''(1)$ を用いると, 系内人数の平均は以下で表せる.

$$\begin{aligned} \mathbb{E}[N] = & \frac{\Pi_0''(1)\sigma}{\lambda_1 + \mu_2} + \frac{\lambda_2^3}{\mu_1 \mu_2 (\lambda_1 + \mu_2)} + \frac{\lambda_2^2(\lambda_1 + \mu_1 + \mu_2)}{\mu_1 \mu_2 (\lambda_1 + \mu_2)} \\ & + \lambda_2 \left\{ \frac{1}{\sigma} + \frac{\lambda_1^2}{\sigma \mu_1 \mu_2} + \frac{\Pi_0''(1)\sigma \mu_2}{\mu_1 (\lambda_1 + \mu_2)^2} + \lambda_1 \left(\frac{1}{\mu_1 (\lambda_1 + \mu_1)} + \frac{1}{\sigma \mu_2} + \frac{1}{\sigma \mu_1} + \frac{\Pi_0''(1)\sigma}{\mu_1 (\lambda_1 + \mu_2)^2} \right) \right\}. \end{aligned}$$

(18), (19) より, 通信が行われていないときの orbit 内の SU の系内人数の母関数は以下となる.

$$\Pi_0(z) = \Pi_0(1) \left(\frac{z_1 - 1}{z_1 - z} \right)^A \left(\frac{z_2 - 1}{z_2 - z} \right)^B. \quad (24)$$

(9), (10) 及び (24) より, 各チャンネルの状態における系内人数の母関数を $\lambda_1, \lambda_2, \mu_1, \mu_2, \sigma$ を用いて表すことができる. (9) 及び (10) より, 次が得られる.

$$\Pi_1''(1) = \lambda_1 \left\{ \frac{\Pi_0''(1) + 2\Pi_2'(1) + \Pi_2''(1)}{\mu_1} + \frac{2\lambda_2(\Pi_0'(1) + \Pi_2(1) + \Pi_2'(1))}{\mu_1^2} + \frac{2\lambda_2^2(\Pi_0(1) + \Pi_2(1))}{\mu_1^3} \right\}, \quad (25)$$

$$\Pi_2''(1) = \frac{\sigma(\gamma''(1)\Pi_0(1) + 2\gamma'(1)\Pi_0'(1) + \gamma(1)\Pi_0''(1))}{\mu_1} + \frac{\lambda_2(3 - \sigma)(\gamma'(1)\Pi_0(1) + \gamma(1)\Pi_0'(1))}{(\lambda_1 + \mu_2)^2} + \frac{2\lambda_2^2\Pi_0(1)}{(\lambda_1 + \mu_2)^3}. \quad (26)$$

よって, (23), (25), (26) 及び $\mathbb{E}[N]$ の結果より, 系内人数の分散を求めることができる. 詳細な結果に関しては省略する.

5 導出結果の考察

本節では, PU・SU の到着率や SU の再試行率を変化させたときに, 前節で求めた安定条件や系内人数の平均・分散がどのように変化するかを考察する. 一般的には, コグニティブ無線では, PU の通信時間よりも SU の通信時間の方が短いことが予想される. そのため, サービス率については, $\mu_1 = 4, \mu_2 = 20$ と設定した. 以下で今回導出した結果を考察していく.

まず, PU と SU の到着率を同時に変化させたときの安定領域について考察していく. 図 2 より, λ_1 の増加に伴い λ_2 が減少していくことが分かる. このことは PU の到着回数が増加することにより, SU の orbit 内の人数が発散しないためには SU の到着回数が減少しなければならないことから明らかである.

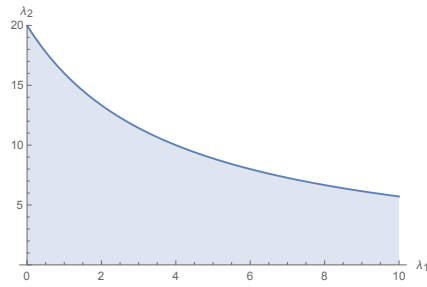


図2 $\mu_1 = 4, \mu_2 = 20$ で, PU と SU の到着率を変化させたときの安定領域. 着色部分が安定領域.

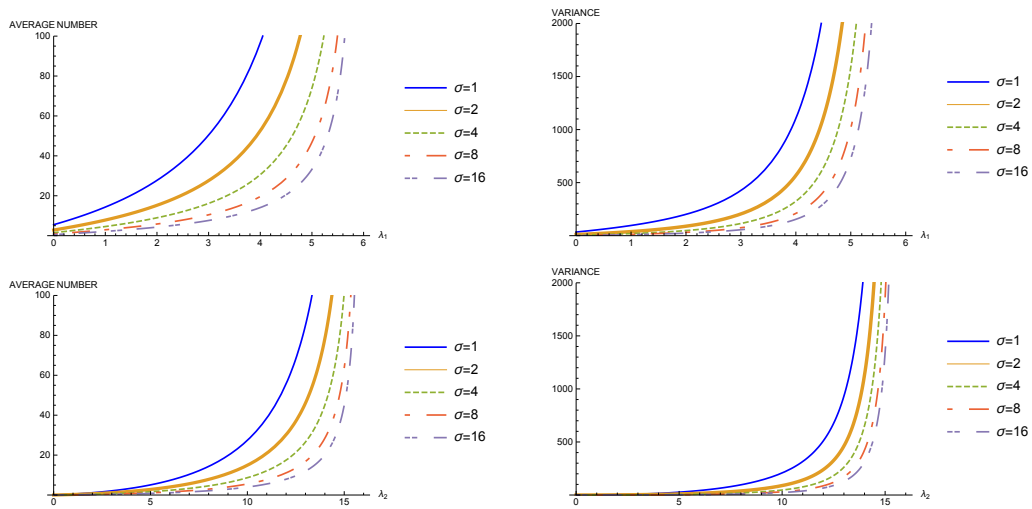


図3 左上図: PU の到着率を変化させたときの orbit 内の SU の平均系内人数の変化. 右上図: PU の到着率を変化させたときの orbit 内の SU の系内人数の分散の変化. 左下図: SU の到着率を変化させたときの orbit 内の SU の平均系内人数の変化. 右下図: SU の到着率を変化させたときの orbit 内の SU の系内人数の分散の変化. 各図において, 実線は $\sigma = 1$, 実太線は $\sigma = 2$, 破線は $\sigma = 4$, 一点鎖線は $\sigma = 8$, 二点鎖線は $\sigma = 16$ を表す.

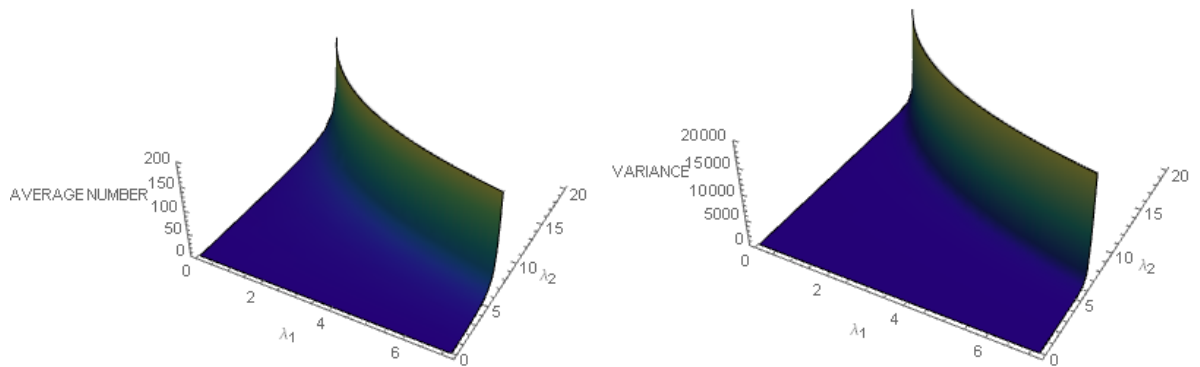


図4 左図: PU および SU の到着率を変化させたときの orbit 内の SU の系内人数の平均の変化. 右図: PU および SU の到着率を変化させたときの orbit 内の SU の系内人数の分散の変化.

次に、安定条件を満たしているときに、PU の到着率を変化させたときの系内人数の平均及び分散について見ていく。ただし、ここでは、 $\lambda_2 = 8$ とし、再試行率は $\sigma = 1, 2, 4, 8, 16$ の場合について考察した。図 3 から読み取れる通り、再試行率が高いほど平均・分散ともに減少していくことが分かる。これは、再試行率が高いほど、orbit 内の各 SU は積極的に再試行しようとするため、その分だけ通信が行われていないときに SU はチャンネルに到達することが多くなる。結果として orbit 内の SU の人数は減り、orbit 内の平均人数も低くなる。さらに orbit 内の SU が少人数で安定するため、ばらつきも少なくなるということから考えて、自然な結果である。

さらに、SU の到着率を変化させたときの系内人数の平均及び分散について見ていく。ただし、 $\lambda_1 = 1$ 、再試行率は $\sigma = 1, 2, 4, 8, 16$ の場合について考察した。こちらも PU の到着率を変化させたときと同様に、再試行率が高いほど平均・分散ともに減少していくことが分かる。

最後に、PU と SU の到着率を同時に変化させたときの系内人数の平均及び分散について考察する。ここでは $\sigma = 8$ とおいた。図 4 より、 λ_1, λ_2 が安定条件に近づくと、orbit 内の SU の系内人数の平均及び分散は急激に上昇することが分かる。これは図 3 の考察の結果とも一致する結果である。

6 おわりに

本稿では、単一チャンネル再試行型待ち行列モデルを用いてコグニティブ無線のモデル化を行った。今回扱ったモデルでは、ユーザが PU と SU という 2 種類のクラスのユーザが存在するものとして、これらのユーザが異なる方法でチャンネルに入り、サービスを受けるという仮定の下で解析をした。このモデルを解析することにより、本稿のモデルの安定条件や SU の orbit 内の平均系内人数及び分散、さらに各サーバの状態における系内人数の母関数を得ることができた。これらの導出した指標を用いて、安定領域や再試行率が異なる場合でのグラフの比較、さらには PU と SU の到着率を同時に変化させたときの平均及び分散の変化についての考察を行った。

今後の課題としては、[12] で導出されているように、SU の再試行回数に関する指標の導出である。再試行型待ち行列モデルにおいて、再試行回数を測定することは待ち時間の測定と同様に、最も重要となる指標の 1 つである。この導出を行うことにより、モデルとしての性能評価を、より具体的に評価できる。また、今回のモデルでは、単一チャンネルと仮定しているが、コグニティブ無線をモデル化するには、様々な周波数帯が存在することから、複数チャンネルを持つモデルが想定される。このような拡張を行うことにより、より現実的なモデルへ近づけていくことが今後の課題である。

参考文献

- [1] 「コグニティブ無線とソフトウェア無線」 <http://sndst.snd.elec.keio.ac.jp/sndgw/RESEARCH/research.files/introduce.pdf> (2018 年 11 月 21 日アクセス)
- [2] 「コグニティブ無線・ホワイトスペースの有効通信利用 (月刊 EMC)」 https://cend.jp/user/headline_technology/20180319.html (2018 年 11 月 30 日アクセス)
- [3] 「コグニティブ無線とは何か？真のコビキタスの実現に必要な不可欠な技術」 <https://www.>

- sbbbit.jp/article/cont1/19899 (2018 年 11 月 30 日アクセス)
- [4] 「ホワイトスペースなど新たな電波の有効利用の実現方策（概要）-総務省」 http://www.soumu.go.jp/main_content/000076997.pdf (2018 年 11 月 21 日アクセス)
 - [5] 「コグニティブ無線 | 用語集 | KDDI 株式会社」 <http://www.kddi.com/yogo/%E9%80%9A%E4%BF%A1%E3%82%B5%E3%83%BC%E3%83%93%E3%82%B9/%E3%82%B3%E3%82%B0%E3%83%8B%E3%83%86%E3%82%A3%E3%83%96%E7%84%A1%E7%B7%9A.html> (2018 年 11 月 30 日アクセス)
 - [6] Y. Konishi, H. Masuyama, S. Kasahara and Y. Takahashi (2013) Performance analysis of dynamic spectrum handoff scheme with variable bandwidth demand of secondary users for cognitive radio networks. *Wireless Networks*, 19, 607-617
 - [7] O. Salameh, K. D. Turck, H. Bruneel, C. Blondia and S. Wittevrongel (2017) Analysis of secondary user performance in cognitive radio networks with reactive spectrum handoff. *Telecommunication Systems*, 65, 539-550
 - [8] J. Keilson, J. Cozzolino and H. Young (1968) A service system with unfilled requests repeated. *Operations Research*, 6, 1126-1137
 - [9] G. I. Falin (1976) Aggregate arrival of customers in a one-line system with repeated calls. *Ukrainian Mathematical Journal*, 28, 561-565
 - [10] G. I. Falin (1977) Waiting time in a single-channel queueing system with repeated calls. *Moscow University Computational Mathematics and Cybernetics*, 4, 83-87
 - [11] G. I. Falin (1980) An M/G/1 system with repeated calls in heavy traffic. *Moscow University Computational Mathematics and Cybernetics*, 6, 48-50
 - [12] G. I. Falin (1986) On the waiting-time process in a single-line queue with repeated calls. *Journal of Applied Probability*, 23, 185-192

移動体通信における Time-based なハンドオフ制限のモデルとその解析

徳山 喜一 三好 直人
東京工業大学 情報理工学院

概要：無線通信セルラネットワークでは、移動体の行うハンドオフ処理の過剰化が処理の失敗等を誘発し、問題となっている。そこで本論文では、移動体のハンドオフレートの低減を目的として、ハンドオフの制限を考えたモデルの下での通信性能の解析を行う。特に、新たなハンドオフ制限のモデルとして、ハンドオフを制限する時間を自由に調整可能なモデルを提案し、当該モデルの下で移動体の通信性能を解析する手法も提案する。解析手法の評価は、数値計算によって行われる。

1. はじめに

1.1. 背景と研究の目的

将来の無線通信セルラネットワークにおいては、日々膨大化し続ける通信の需要に対応するため、通信を供給する基地局の負荷分散を目的として基地局をさらに高密度に配置する検討がなされている。その一方で、高密度化されたセルラネットワークではユーザーの行うハンドオフの過剰化が深刻な問題となる。ハンドオフとは、ユーザー（移動体）と通信先の基地局との距離が遠ざかった場合に、別のより近い基地局へ通信先を変更する処理のことであり、特に高密度な基地局の間を移動体が高速で移動する場合に、ハンドオフの過剰化による通信の不安定化やハンドオフの処理の失敗等のリスクが発生する。

高密度化されたセルラネットワークにおける移動体のハンドオフの過剰化を防ぐため、移動体の速度に応じてハンドオフを制限する手法がこれまでに提案されており、数理モデル上の評価によってその有用性が確認されている [1] [2]。ハンドオフの制限とはすなわち、つねに最も近い基地局との通信を保つようなハンドオフを移動体は必ずしも行わないという通信のモデルである。一方で、移動体のハンドオフの制限は移動体と通信先の基地局との距離の増大、すなわち受信データ量の減少を引き起こす。よって、移動体に対するハンドオフの制限と受信データ量とは互いにトレードオフの関係にあるため、ハンドオフの制限はその強度が適度に調整される必要がある。

本研究では、セルラネットワーク上でハンドオフの制限を行う移動体に対し、その通信性能を解析可能な新たなモデルを提案する。特に、ハンドオフの制限の強度が任意に調整可能である場合に、ハンドオフの制限を行う移動体の通信性能を、受信データ量とハンドオフレートのトレードオフによって評価する。さらに、移動体の移動速度や基地局の配置密度等に応じて、移動体の通信性能を最大化させる最適なハンドオフの制限強度の導出も試みる。

1.2. 関連研究

確率幾何モデルの枠組みにおいてハンドオフ制限のモデルの提案とその評価を行った例として、まず Arshad らの研究 [1] が挙げられる。当該研究では、Single-tier のセルラネットワーク上でユーザー端末が任意の曲線軌道上を移動する場合を考え、ユーザー端末のハンドオフレートと平均受信データレートの解析が行われた。特に、ユーザー端末のハンドオフに関して、ユーザー端末の通信先である基地局が自身から 1 番目、もしくは 2 番目に近い基地局である間はユーザー端末はハンドオフを行わないというポリシーが定められた。これは、ユーザー端末が常に自身から最短距離にある基地局と通信をするようハンドオフを行うという通常のポリシーに対し、ユーザー端末のハンドオフレートが小さくなるが、一方でユーザー端末の平均受信データレートも小さくなる。このようなハンドオフを制限するというポリシーの下で、当該研究ではユーザー端末のハンドオフレートと平均受信データレートを確率幾何モデル上で計算し、両者のトレードオフ関係を表した評価関数によってその通信性能を評価した。その結果、当該研究で定めたハンドオフ制限のポリシーでは、ハンドオフ制限を行わない場合に対して通信性能が上回る場合があることが観察された。特に、基地局の密度が高く、かつユーザー端末の移動速度も高い場合に、ハンドオフ制限のポリシーは通信性能において有利になることが確認された。なお、Arshad らの別の研究結果 [2] において、ユーザー端末の通信先である基地局が自

身から4番目に近い基地局である間はハンドオフを行わないという新たなハンドオフ制限に関するポリシーを提案し、同様の評価を行っている。

これらの研究において考えられたハンドオフ制限のモデルでは、ユーザー端末の通信先である基地局が自身から何番目に近い基地局であるかに着目しており、これらのモデルを Number-based なハンドオフ制限のモデルと呼ぶことにする。これに対し、本研究では、移動するユーザー端末が最後にハンドオフを行った時点からの経過時間の制御によってハンドオフ制限の程度を決定するという、Time-based なハンドオフ制限のモデルを提案する。

1.3. 本研究の成果

本研究では、基地局の配置が一様なポワソン点過程によって配置されるという Single-tier のセルラネットワークモデルを考え、上述した Time-based なハンドオフ制限のモデルについて定義を与える。さらに、定義された Time-based なハンドオフ制限モデルの下で、ユーザー端末のハンドオフレートと平均受信データレートの解析的な計算結果を与える。特に、Time-based なハンドオフ制限のモデルではユーザー端末がハンドオフを制限する時間 s がパラメータとして与えられるため、ハンドオフレートと平均受信データレートはハンドオフ制限時間 s をパラメータとする関数となることに注意する。そして、Time-based なハンドオフ制限モデルにおけるユーザー端末のスループット関数を、ハンドオフレートと平均受信データレートのトレードオフを表すよう定義し、パラメータ s に関するスループット関数の推移を観察する。その結果、一定条件下でスループット関数は s に関する最大値をもつことを数値計算により確認した。すなわち、一定のハンドオフ制限時間を設けることでユーザー端末のスループットが増加しうることが本研究によって示された。さらに本研究では、スループット関数がハンドオフ制限時間 s に関して最大値を持つための必要十分条件や、これらの最大値を与えるパラメータ s の値 s^* を解析的に計算するための手法についても考察した。

2. システムモデル

2.1. ネットワークモデル

単一周波数帯の通信のみを扱うダウンリンク型の無線通信ネットワーク（ホモジニアスネットワーク）を考える。ここで、通信を供給する基地局は配置強度を $\lambda (\in (0, \infty))$ に持つ $\mathbb{R}^2$ 上の定常ポワソン点過程 $\Phi = \{X_1, X_2, \dots\}$ に従って配置されるとし、各基地局から発信される電波は全て強度が同一の1であるとする。通信端末が基地局から電波を受信する際、その電波の強度は $\beta \geq 2$ の指数係数によって減衰するとする。また、受信電波に係るフェーディングは Rayleigh Fading を仮定するが、シャドローイングは無視するものとする。すなわち、時刻 t において地点 $\mathbf{u} \in \mathbb{R}^2$ に位置するユーザー端末が、地点 $X_i \in \mathbb{R}^2$ に配置された基地局から受信する電波の強度は $H_{i,t} |\mathbf{X}_i - \mathbf{u}|^{-\beta}$ と表される。ここで、 $H_{i,t}$ は $i \in \mathbb{N}$ 、 $t \in \mathbb{N}_0 = \{0\} \cup \mathbb{N}$ に関して互いに独立で、かつ平均1の指数分布に従う確率変数である。

ここで、各ユーザー端末が自身から最も近い基地局と通信を行うことを仮定する場合、各基地局の通信区画は図.1 に示されるようなポワソン-ボロノイ分割と同一になる。このとき、原点に配置されたユーザー端末と通信先である基地局を X_{b_o} と表すとき、 $R_{b_o} = |X_{b_o}|$ の密度関数はよく知られており ([3], Section 2.3)、以下のように与えられる。

$$f_{b_o}(r) = 2\pi\lambda r e^{-\lambda\pi r^2}, \quad r \geq 0.$$

さらに、ユーザー端末は2次元平面上を移動しており、基地局は単位時間ごとに、各々の通信区画内に配置されたユーザー端末へ通信電波を供給する。時刻 $t \in \mathbb{N}_0$ において地点 $\mathbf{u} \in \mathbb{R}^2$ に位置するユーザー端末が、地点 $X_i \in \mathbb{R}^2$ に位置する基地局と通信を行うとき、ダウンリンク通信における信号対干渉比 (Signal-to-Interference Ratio, SIR) は以下のように与えられる。

$$\text{SIR}_{u,i} = \frac{H_{i,t} |\mathbf{X}_i - \mathbf{u}|^{-\beta}}{I_{u,i}}.$$

ここで、 $I_{u,i}$ はユーザー端末に対する干渉電波の強さを表し、以下のように定式化される。

$$I_{u,i} = \sum_{j \in \mathbb{N} \setminus \{i\}} \frac{H_{j,t}}{|\mathbf{X}_j - \mathbf{u}|^\beta}. \quad (1)$$

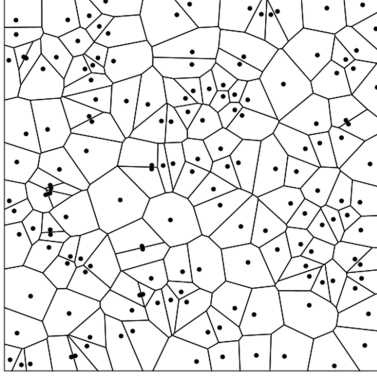


図 1: ポワソン-ボロノイ分割のサンプル

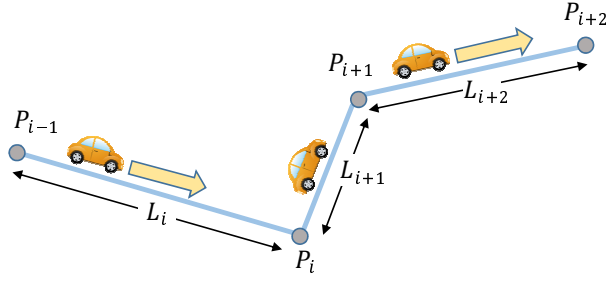


図 2: ユーザー端末の移動軌道の例

以上を用いて、単位時間における平均受信データレートは以下のように定義される。

$$\tau_{u,i}(\lambda, \beta) = E[\log(1 + \text{SIR}_{u,i})]. \quad (2)$$

2.2. ユーザー端末の移動軌道モデル

本研究では、移動体ユーザー端末の移動モデルとして、以下に定めるようなランダムウォークモデルを仮定する。すなわち、各ユーザーは時間 s の間、任意方向に一直線上を移動するものとし、時間 s の期間を 1 つの移動ピリオドと定める。第 i 番目の移動ピリオドにおけるユーザー端末の移動速度を v_i とし、特に 1 つの移動ピリオドにおける移動速度は一定とする。ユーザー端末は 1 つの移動ピリオドが終了した時点で任意に方向を変えてよいものとする。第 i 番目の移動ピリオドの始点を $P_i \in \mathbb{R}^2$ 、終点を $P_{i+1} \in \mathbb{R}^2$ 、移動ピリオドの間にユーザー端末が移動する距離を $L_i = |P_{i+1} - P_i|$ と表す ($L_i = sv_i$)。図.2 はユーザー端末の移動モデルの例である。特に、各移動ピリオドにおいてユーザー端末が方向転換を行わない場合、ユーザー端末の移動軌道は直線になることに注意する。

2.3. Time-based なハンドオフ制限のモデル

ここでは、ユーザー端末に対しハンドオフを制限させるためのあるポリシーを定め、これをユーザー端末に対するハンドオフ制限のモデルとして定義する。特に、ユーザー端末のハンドオフを一定時間制限するというハンドオフ制限のモデルをここでは定義する。すなわち、先の節で定義したユーザー端末の移動軌道モデルにおいて、ユーザー端末は各地点 P_i においてのみハンドオフを行い、 P_i から P_{i+1} へ移動する間はハンドオフ処理を一切行わないことを仮定する。このとき、ユーザー端末は時間 s が経過するごとにハンドオフを行うこととなり、これを時間 s をハンドオフ制限強度とする Time-based なハンドオフ制限モデルと定める。ここで、地点 P_{i+1} において実際にハンドオフが行われるための条件は、 P_i 、 P_{i+1} からそれぞれ最短距離の基地局が互いに異なっている場合であることに注意し、この事象の確率を $H(s, v_i, \lambda)$ と定める。また、ユーザー端末が地点 P_i から距離 u ($0 \leq u \leq L_i = sv_i$) だけ移動した地点における平均受信データレートを $\tau(u, \lambda, \beta)$ と定める。

図.3 の左側の結果は、 $\tau(u, \lambda, \beta)$ の u 推移をシミュレーションにより与えたものである ($(s, v_i, \lambda, \beta) = (15, 0.1, 4, 4)$)。以上の結果から分かるように、 $\tau(u, \lambda, \beta)$ は u に関する減少関数となる。これは、ハンドオフが行われる地点 P_i においてはユーザー端末が最短距離の基地局と通信を行う一方で、ユーザー端末が地点 P_i から離れるに従い、通信先である基地局とユーザー端末との平均距離が増大していくことが理由である。なお、地点 P_{i+1} に到達した時点でユーザーはハンドオフを行うため、 P_{i+1} における平均受信データレートは P_i におけるものと等しくなる ($\tau(0, \lambda, \beta) = \tau(sv_i, \lambda, \beta)$)。次に、図.3 の右側の結果は $H(s, v_i, \lambda)$ の s 推移をシミュレーションにより与えたものである ($(s, v_i, \lambda) = (15, 0.2, 4)$)。以上の結果から分かるように、 v_i が一定であるとき、 $H(s, v_i, \lambda)$ は s に関する増加関数である。すなわち、ユーザー端末の移動距離 L_i が増加するに従い、地点 P_{i+1} においてハンドオフが起こる確率は最大値の 1 まで増大を続ける。

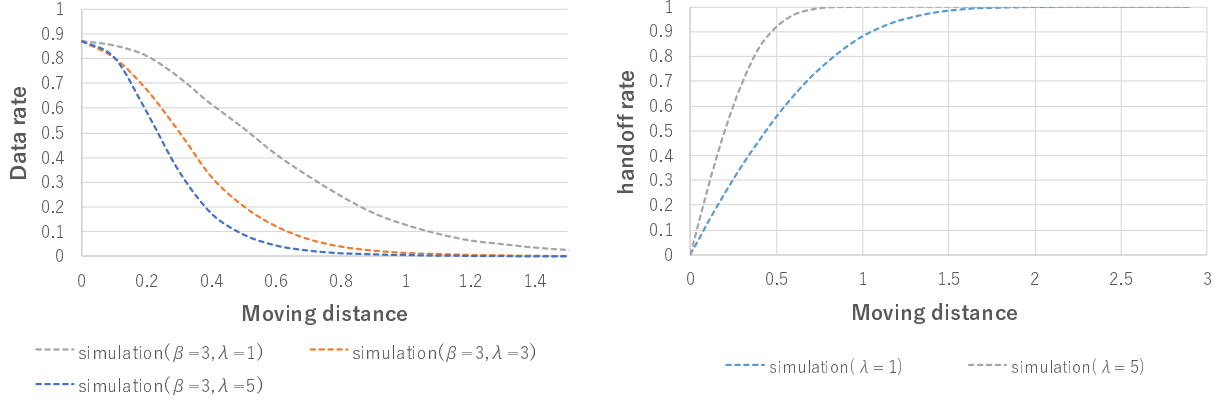


図 3: (左): $\tau(u, \lambda, \beta)$ のシミュレーション計算の数値例
(右): $H(s, v_i, \lambda)$ のシミュレーション計算の数値例

2.4. ユーザー端末のスループット

Time-based なハンドオフ制限のモデルでは、ハンドオフ制限強度 s の調整により、1 つの移動ピリオドの間にユーザー端末が移動する距離が変化する。これによりユーザー端末がハンドオフを制限しながら移動する距離が変化するため、ユーザー端末のハンドオフレートをパラメータ s により調整することができる。導入で述べたように、ハンドオフレートと平均受信データレートとは互いにトレードオフの関係にあるため、制限強度パラメータ s を調整することでトレードオフ関係を最適化することが肝要となる。以上のことを目的としてユーザー端末のスループットを表す評価関数 $Q_{v_i, \lambda, \beta}(s)$ を定める。特に、関連研究 [4], [5] にならい、評価関数 $Q_{v_i, \lambda, \beta}(s)$ を以下のように定義する。

$$Q_{v_i, \lambda, \beta}(s) = \frac{1}{s} \{ D(s, v_i, \lambda, \beta) - C \cdot H(s, v_i, \lambda) \}. \quad (3)$$

ここで、 $D(s, v_i, \lambda, \beta)$ は第 i 番目の移動ピリオド全体におけるユーザー端末の送受信データ量の平均である。特に、 $\frac{1}{s} D(s, v_i, \lambda, \beta)$ 、 $\frac{1}{s} H(s, v_i, \lambda)$ はそれぞれ、Time-based なハンドオフ制限モデルにおけるハンドオフレートと平均受信データレートを表すことに注意する。 v_i, λ, β に対し、 $Q_{v_i, \lambda, \beta}(s)$ は s の関数として与えられることが分かる。

3. ハンドオフレートと平均受信データレートの解析

Time-based なハンドオフ制限モデルにおけるハンドオフレート $\frac{1}{s} D(s, v_i, \lambda, \beta)$ と平均受信データレート $\frac{1}{s} H(s, v_i, \lambda)$ について解析的な計算結果を得ることができれば、式 (3) により与えられたスループット関数の評価が可能となる。この章では、式 (3) で与えられた $D(s, v_i, \lambda, \beta)$ と $H(s, v_i, \lambda)$ をそれぞれ解析的に計算するための手法を提案する。以後、ユーザー端末の移動軌道において特定の 1 つの移動ピリオドに着目し、ユーザー端末の出発点を $\mathbf{o} = (0, 0) \in \mathbb{R}^2$ 、終点を $(sv, 0) \in \mathbb{R}^2$ と定める。特にこれらの仮定は、基地局の配置を表すポワソン点過程 Φ が定常性を持つという前提から、一般性を失わずに定められる。

3.1. 解析結果

本論文において定義されたモデルにおけるハンドオフレートと平均受信データレートの解析結果は、筆者の過去の論文においてそれぞれ結果が与えられている [6], [7]。ここでは [7] においてまとめた結果を引用して述べる。

定理 1 ダウンリンク通信のホモジニアスセルラネットワークにおいて、ユーザー端末の移動軌道は第 2.2 節で定められたものとし、またユーザー端末のハンドオフに関するポリシーは第 2.3 節で定められた Time-based なハンドオフ制限モデルに従うものとする。このとき、ユーザー端末のハンドオ

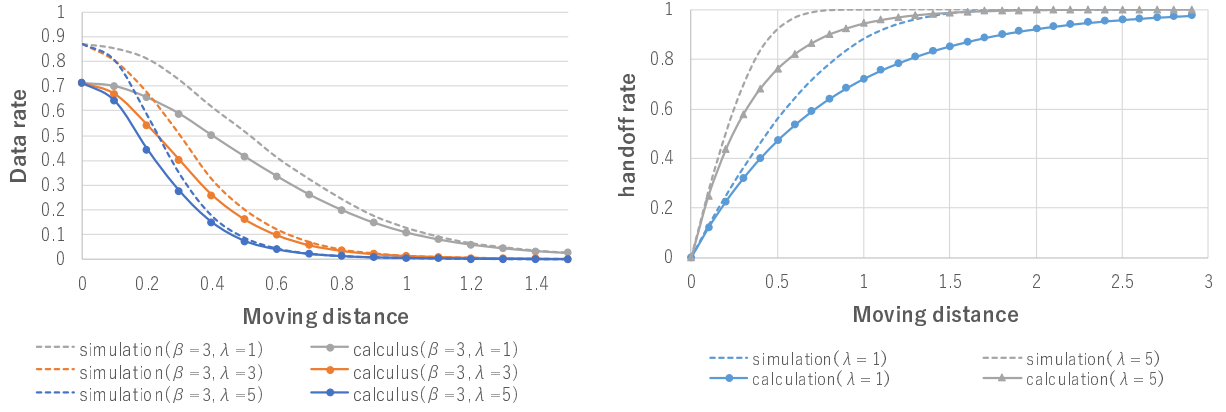


図 4: (左): $\tau(u, \lambda, \beta)$ の解析計算/シミュレーション計算の比較例
(右): $H(s, v, \lambda)$ の解析計算/シミュレーション計算の比較例

フレート $\frac{1}{s}H(s, v, \lambda)$ と平均受信データレート $\frac{1}{s}D(s, v, \lambda, \beta)$ は近似的に以下のように与えられる。

$$\frac{1}{s}H(s, v, \lambda) \approx \frac{1}{s} \left\{ 1 - e^{-\frac{4sv\sqrt{\lambda}}{\pi}} \right\}. \quad (4)$$

$$\frac{1}{s}D(s, v, \lambda, \beta) \approx \frac{1}{v} \int_0^v \tau(sw, \lambda, \beta) dw, \quad (5)$$

ここで、 $\tau(u, \lambda, \beta)$ は以下のように近似的に表される。

$$\tau(u, \lambda, \beta) = \int_0^\infty \frac{1}{1+z} \cdot \frac{\exp\left(-\frac{\pi\lambda}{1+K_{z,\beta}^{-1}}u^2\right)}{1+K_{z,\beta}} dz, \quad (6)$$

$$K_{z,\beta} = \frac{2\pi}{\beta} z^{\frac{2}{\beta}} \csc \frac{2\pi}{\beta}. \quad (7)$$

注意 1 セルラネットワークにおけるハンドオフレートは、基地局のセル区画を表すポワソン-ボロノイ分割について、その境界辺とユーザー端末の移動軌道との交点の数によって計算可能である。ここで式 (4) は、ポワソン-ボロノイ分割の境界辺の中点がポワソン点過程によって配置されるという近似を導入し、得られた結果である。また、式 (6) は式 (2) で定義された平均受信データレートの計算により得られるが、特に、干渉電波を発する基地局が $\mathbb{R}^2$ 全体に配置されると仮定したこと、すなわち式 (1) を以下のように近似したことにより得られた。

$$I_{u,i} = \sum_{j \in \mathbb{N} \setminus \{i\}} \frac{H_{j,t}}{|X_j - u|^\beta} \approx \sum_{j \in \mathbb{N}} \frac{H_{j,t}}{|X_j - u|^\beta}. \quad (8)$$

3.2. シミュレーションとの比較結果

図.4 は、式 (4),(6) においてそれぞれ得られた $H(s, v, \lambda)$ 、 $\tau(u, \lambda, \beta)$ について、それぞれシミュレーションとの比較を行った結果である ($0 \leq u \leq sv$)。双方の結果において、シミュレーションとの値に差異が見られることが分かる。

ハンドオフ確率 $H(s, v, \lambda)$ の結果については、特に解析結果はシミュレーションの結果に対し値が小さくなっていることが分かる。解析計算においては、ポワソン-ボロノイ分割を構成する境界辺の中点がポワソン点過程に従うことを仮定しているが、本来の境界辺の中点が従う分布はポワソン点過程に対して点のばらつきが起りにくいと考えられる。このため、距離 sv の直線軌道がポワソン-ボロノイ分割の境界辺を全く交点を持たない確率 ($= 1 - H(s, v, \lambda)$) は解析結果の方が大きくなり、その結果ハンドオフが起こる確率 $H(s, v, \lambda)$ は解析結果の方が下回ってしまうと考えられる。これらの誤差の改善のためには、ポワソン-ボロノイ分割の境界辺の中点の従う分布をより正確に定めることが重要である。

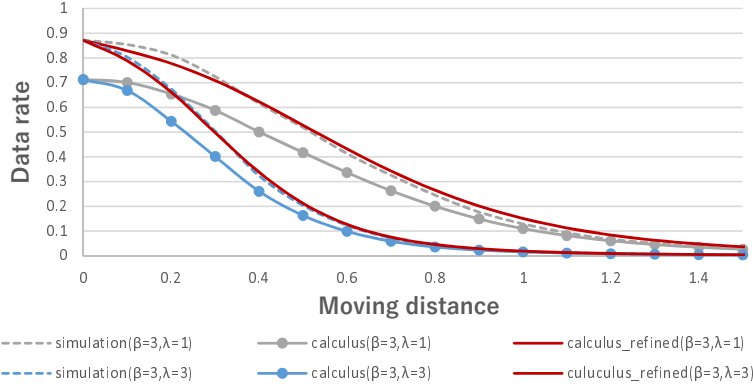


図 5: $\tau_2(u, \lambda, \beta)$ の解析計算結果の数値例

地点 $(u, 0)$ における平均受信データレート $\tau(u, \lambda, \beta)$ の結果についても、シミュレーションの結果に対して解析結果は値が小さくなっていることが分かる。特に、原点 $\mathbf{o}$ からの移動距離 u が小さい場合に解析結果とシミュレーションとの誤差が大きくなり、逆に移動距離 u が大きくなるにつれてその誤差が小さくなるが、これは式 (8) において導入された近似の影響である。ユーザー端末は原点 $\mathbf{o}$ から最も近い基地局 X_{b_o} と通信を行うため、干渉になる基地局は本来、原点を中心とする半径 $|X_{b_o}|$ の円 $B(\mathbf{o}, |X_{b_o}|)$ の内部には位置しない。よって、式 (8) の近似により干渉になる基地局が $\mathbb{R}^2$ 上全体に配置される場合、 $B(\mathbf{o}, |X_{b_o}|)$ の内部にも干渉の基地局が配置されてしまう影響により原点 $\mathbf{o}$ 付近で近似の影響が強く現れてしまうことが、これらの誤差の原因と考えられる。

3.3. 解析結果の改良

ここでは、先ほどの評価結果でシミュレーション結果との誤差が確認された解析結果に対し、誤差を改善するための手法を提案する。特に、地点 $(u, 0)$ における平均受信データレート $\tau(u, \lambda, \beta)$ の解析結果に見られる誤差の改善方法について考察する。

$\tau(u, \lambda, \beta)$ の解析結果では、特に u の値が小さい場合に誤差が大きく観測された。一方で、 $u = 0$ の場合の平均受信データレート $\tau_0(\lambda, \beta) = \tau(0, \lambda, \beta)$ を近似なしで計算するための手法は [8] によって提案されており、以下のように与えられる ([8], Theorem3)。

$$\tau_0(\lambda, \beta) = \int_0^\infty \int_0^\infty \frac{\rho_1(z, u, \lambda, \beta)}{1+z} dz du,$$

ただし

$$\rho_1(z, u, \lambda, \beta) = \exp \left\{ -\sigma^2 z \left(\frac{u}{\pi \lambda} \right)^{\beta/2} - u \left(1 + \frac{2z^{2/\beta}}{\beta} \int_{1/z}^\infty \frac{v^{2/\beta-1}}{1+v} dv \right) \right\}.$$

以上を用いて、ハンドオフの制限を仮定した場合の地点 $(u, 0)$ における平均受信データレート $\tau_2(u, \lambda, \beta)$ を新たに以下のように定義する。

$$\tau_2(u, \lambda, \beta) = e^{-\nu u} \tau_0(\lambda, \beta) + (1 - e^{-\nu u}) \tau(u, \lambda, \beta), \quad 0 \leq u \leq sv.$$

ここで、 ν は u に関する重みづけパラメータである。すなわち、原点からの移動距離 u の増加にしたがつて $\tau_0(\lambda, \beta)$ に寄与する重みが小さくなると同時に、 $\tau(u, \lambda, \beta)$ に寄与する重みが大きくなる。 $\tau(u, \lambda, \beta)$ は特に u が小さい場合にシミュレーション結果との誤差が大きく現れることから、これらの重み付けによって誤差を修正可能である。

図.5 は、Time-base なハンドオフ制限モデルにおける地点 $(u, 0)$ での平均受信データレートについて、シミュレーション結果に対する解析結果 $\tau(u, \lambda, \beta)$ と解析結果 $\tau_2(u, \lambda, \beta)$ のプロットである ($0 \leq u \leq sv$)。ここで、 $\tau_2(u, \lambda, \beta)$ におけるパラメータ ν はシミュレーション結果に対する誤差の最小二乗法により値を決定した。以上の結果より、改良した解析結果である $\tau_2(u, \lambda, \beta)$ では、適切なパラメータ ν を決定することでシミュレーション結果との誤差が小さい値を導出可能であることが分かる。ここで、最小二乗法により決定される ν の値は、パラメータ λ, β の値によって変化することに注意する。すなわち、 λ, β の値に対してパラメータ ν の最適値の推定を都度行う必要がある。

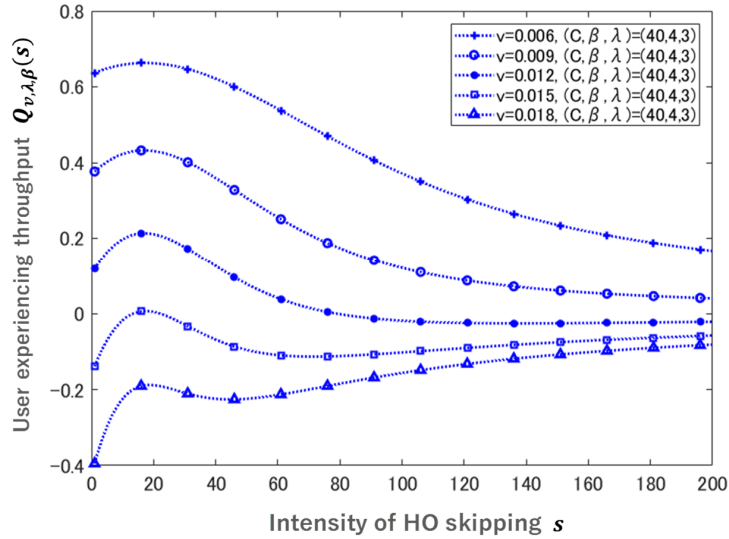


図 6: $Q_{v,\lambda,\beta}(s)$ の s 推移 数値計算例

4. スループット関数の解析と評価

この章では、式 (3) によって定められたユーザー端末のスループットを表す関数 $Q_{v,\lambda,\beta}(s)$ について、数値計算による評価と解析的な評価についていくつかの結果を与える。式 (3) を構成するハンドオフレートと平均受信データレートについては、第 3 章の定理 1 によって解析的な計算の結果がそれぞれ与えられたため、これにより $Q_{v,\lambda,\beta}(s)$ の評価が可能となった。ここではスループット関数 $Q_{v,\lambda,\beta}(s)$ のハンドオフ制限強度 s に関する推移を観察し、特に、 $Q_{v,\lambda,\beta}(s)$ はある一定の条件の下で s に関する最大値を持つことを確認する。その上で、 $Q_{v,\lambda,\beta}(s)$ が最大値を持つための必要十分条件について考察する。最後に、 $Q_{v,\lambda,\beta}(s)$ の最大値を与えるパラメータ s の値 s^* についての解析的な導出方法について考察する。

4.1. スループット関数の数値計算評価

スループット関数 $Q_{v,\lambda,\beta}(s)$ の s 推移に関して、数値的な計算結果を図.6 に示す。図.6 では、パラメータ λ, β は固定され、パラメータ v は 5 つの値のパターンで出力されている ($(\lambda, \beta) = (3, 4), v = 0.006, 0.009, 0.012, 0.015, 0.018$)。図.6 の結果より、特に v の値が小さい場合において、スループット関数 $Q_{v,\lambda,\beta}(s)$ は s に関する最大値をもつことが分かる。すなわち、一定の制限強度によってハンドオフの制限を与えた場合にユーザー端末のスループットは最大値を持つということであり、これはハンドオフ制限モデルの有効性を示した結果である。一方で図.6 の結果から、 v の値が一定以上大きい場合にはスループット関数 $Q_{v,\lambda,\beta}(s)$ は最大値を与えないことも同様に分かる。ハンドオフ制限強度 s の増加によって $Q_{v,\lambda,\beta}(s)$ は 0 に収束していくが、ユーザー端末の移動速度 v が十分大きい場合には $Q_{v,\lambda,\beta}(s)$ が常に負の値を取るため、 $Q_{v,\lambda,\beta}(s)$ が s に関して単調増加となってしまうことが、最大値を持たない原因であると考えられる。さらに図.6 から分かる結果として、スループット関数 $Q_{v,\lambda,\beta}(s)$ の最大値を与える s の値は、各 v の値に対してほぼ不変であることが分かる。これらの理由については現時点で明確ではなく、理由の明確化は今後の課題である。

さらに図.7 では、 $Q_{v,\lambda,\beta}(s)$ の最大値を与えるパラメータ s の値 s^* を様々な C, λ, β の値に対して数値計算により出力した結果である。図.7 の結果から、 s^* の値はハンドオフコスト C に関して比例して増加するが、パスロス係数 β に関しては減少する性質を持つことが分かる。さらに図.8 の結果から、 s^* の値は基地局の配置密度 λ に関してほぼ不変であることが分かる。現時点では、これらの結果が与えられる理由についても明らかになっていない。

4.2. スループット関数が最大値を持つ条件

これまでの結果から、式 (3) によって定められたスループット関数 $Q_{v,\lambda,\beta}(s)$ は s に関して最大値を持つ場合があることが確認された。特に、ユーザー端末の移動速度 v の値により、 $Q_{v,\lambda,\beta}(s)$ が最大値を持つかどうかの結果が変わることを確認した。さらに、 $Q_{v,\lambda,\beta}(s)$ が単峰である、すなわち $Q_{v,\lambda,\beta}(s)$ がただ 1 つの極大値を持つための必要十分条件について以下のことが言える。

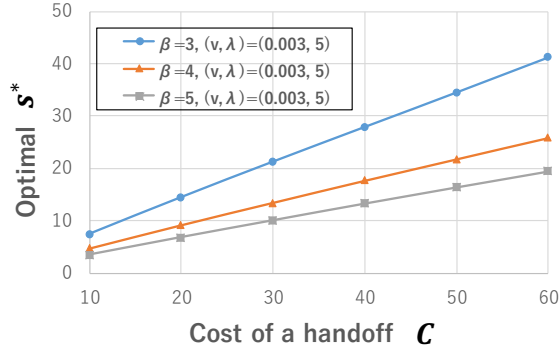


図 7: $\frac{d}{ds}Q_{v,\lambda,\beta}(s)$ の s, v 推移

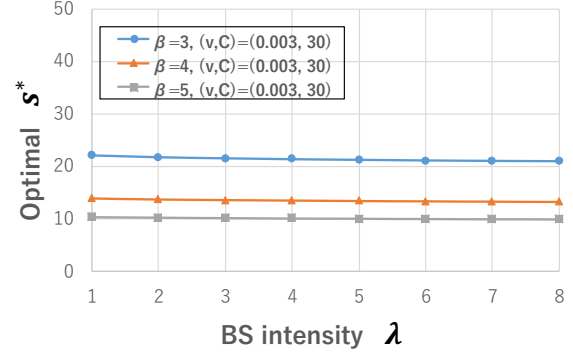


図 8: s^* の解析計算/数値計算結果の比較例

命題 1 以下の事柄が成り立つことを仮定する。

(仮定 A) 任意の $v, \lambda, \beta > 0$ に対し、式 (3) で与えられる $Q_{v,\lambda,\beta}(s)$ は $s \geq 0$ 上で高々 2 つの極値を持つ。

このとき、 $Q_{v,\lambda,\beta}(s)$ が単峰であるための必要十分条件は、以下のように与えられる。

$$v < \frac{1}{2C\sqrt{\lambda}} \int_0^\infty \frac{1}{(1+z)\sqrt{K_{z,\beta}(1+K_{z,\beta})}} dz, \quad (9)$$

ここで $K_{z,\beta}$ は式 (7) で与えられる。

[証明] (仮定 A) が成り立つ場合、 $Q_{v,\lambda,\beta}(s)$ が単峰であるためには (i) $\lim_{s \rightarrow 0} \frac{d}{ds}Q_{v,\lambda,\beta}(s) > 0$ か (ii) $\lim_{s \rightarrow \infty} \frac{d}{ds}Q_{v,\lambda,\beta}(s) < 0$ であることが必要十分である。式 (4),(5) を用いて式 (3) で表される $Q_{v,\lambda,\beta}(s)$ を計算すれば、以下のようになる。

$$\begin{aligned} s^2 \times \frac{d}{ds}Q_{v,\lambda,\beta}(s) &= -\frac{1}{v} \int_0^{sv} \tau(u, \lambda, \beta) du + C(1 - e^{-\frac{4sv\sqrt{\lambda}}{\pi}}) + s \cdot \tau(sv, \lambda, \beta) - C \cdot \frac{4sv\sqrt{\lambda}}{\pi} e^{-\frac{4sv\sqrt{\lambda}}{\pi}} \\ &\xrightarrow{s \rightarrow \infty} -\frac{1}{v} \int_0^\infty \tau(u, \lambda, \beta) du + C < 0. \end{aligned} \quad (10)$$

ここで、式 (10) における最後の不等式に式 (6) を代入すれば、これは式 (9) と同値である。すなわち、(ii) $\lim_{s \rightarrow \infty} \frac{d}{ds}Q_{v,\lambda,\beta}(s) < 0$ と式 (9) は互いに同値である。

あとは (i) $\lim_{s \rightarrow 0} \frac{d}{ds}Q_{v,\lambda,\beta}(s) > 0$ を示せばよい。 s が十分小さい場合、式 (6) で表される $\tau(u, \lambda, \beta)$ に多項式による近似を適用可能であり、 $Q_{v,\lambda,\beta}(s)$ は以下のように表される。

$$\begin{aligned} Q_{v,\lambda,\beta}(s) &\approx \int_0^\infty \frac{1}{(1+z)(1+K_{z,\beta})} \left\{ 1 - \frac{\pi\lambda}{3(1+K_{z,\beta}^{-1})} s^2 v^2 \right\} dz - C \cdot \frac{4v\sqrt{\lambda}}{\pi} + C \cdot \frac{8sv^2\lambda}{\pi^2}. \end{aligned} \quad (11)$$

よって、 $\frac{d}{ds}Q_{v,\lambda,\beta}(s)$ は以下のように計算される。

$$\begin{aligned} \frac{d}{ds}Q_{v,\lambda,\beta}(s) &\approx -\frac{2}{3}\pi\lambda sv^2 \int_0^\infty \frac{K_{z,\beta}}{(1+z)(1+K_{z,\beta})^2} dz + C \cdot \frac{8v^2\lambda}{\pi^2} \\ &\xrightarrow{s \rightarrow 0} C \cdot \frac{8v^2\lambda}{\pi^2} > 0. \end{aligned} \quad (12)$$

以上より命題の結果が従う。 $\square$

命題 1 の (仮定 A) については、数値計算によりその正当性が評価できる。図.9 は $\frac{d}{ds}Q_{v,\lambda,\beta}(s)$ の s, v に関する 3 次元プロット結果であり、図中の灰色の平面は $\frac{d}{ds}Q_{v,\lambda,\beta}(s) = 0$ を表す $((\lambda, \beta) = (3, 4))$ 。図.9 の結果より、 $\frac{d}{ds}Q_{v,\lambda,\beta}(s) = 0$ の s に関する解は高々 2 つであり、すなわち $Q_{v,\lambda,\beta}(s)$ の s に関する極値は高々 2 つであることが予想される。このため、命題 1 の (仮定 A) はみたされ、式 (3) で与えられる $Q_{v,\lambda,\beta}(s)$ に対し (9) の必要十分条件が適用可能であることが期待される。

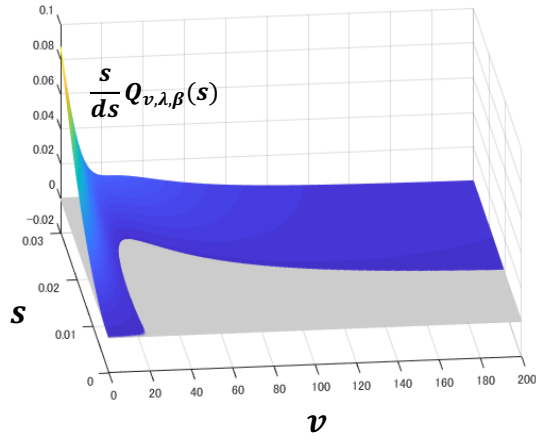


図 9: $\frac{d}{ds} Q_{v,\lambda,\beta}(s)$ の s, v 推移

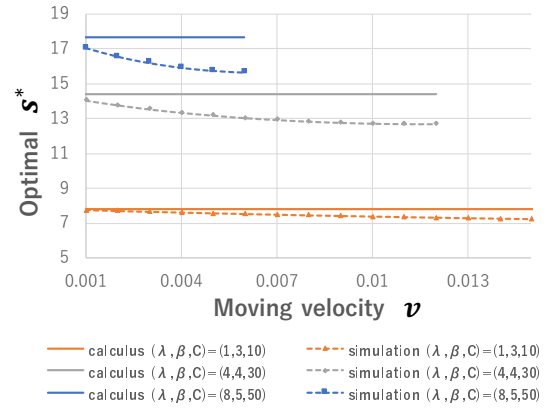


図 10: s^* の解析計算/数値計算結果の比較例

4.3. 最適ハンドオフ制限強度の解析的表現

図.6 の評価結果ではスループット関数 $Q_{v,\lambda,\beta}(s)$ が s についての最大値を与えることを観察した。ここでは、スループット関数 $Q_{v,\lambda,\beta}(s)$ に対して v を漸近的に 0 へ近づけることで、 $Q_{v,\lambda,\beta}(s)$ に多項式近似を適用し、 s^* の値を解析的に導出することを考える。特に図.6 の評価では、ユーザー端末の移動速度 v に関して $Q_{v,\lambda,\beta}(s)$ の最大値を与える s の値 s^* はほとんど変わっていないことを観察した。この性質を利用し、 $v \rightarrow 0$ として求めた s^* の解析的な導出結果と、 s^* の数値計算による導出結果とを比較し、特に v が十分大きい場合に両者の差異を観察する。

この節における主要の結果は以下である。

定理 2 式 (3) で与えられたスループット関数 $Q_{v,\lambda,\beta}(s)$ に対し、命題 1 の (仮定 A) が成り立つものとする。このとき、 $Q_{v,\lambda,\beta}(s)$ の s についての最大値を与える値 s^* について以下が成り立つ。

$$s^* \xrightarrow{v \rightarrow 0} \frac{12C}{\pi^3} \left\{ \int_0^\infty \frac{K_{z,\beta}}{(1+z)(1+K_{z,\beta})^2} dz \right\}^{-1}, \quad (13)$$

ここで $K_{z,\beta}$ は式 (7) で与えられる。

[証明] 式 (3) で与えられた $Q_{v,\lambda,\beta}(s)$ に対し漸近近似 $v \rightarrow 0$ を考えた場合の導関数 $\frac{d}{ds} Q_{v,\lambda,\beta}(s)$ を計算する。計算方法は式 (10), (11), (12) において示したものと同様であり、結果は以下のようになる。

$$\frac{d}{ds} Q_{v,\lambda,\beta}(s) \xrightarrow{v \rightarrow 0} -\frac{2}{3} \pi \lambda s v^2 \int_0^\infty \frac{K_{z,\beta}}{(1+z)(1+K_{z,\beta})^2} dz + C \cdot \frac{8v^2 \lambda}{\pi^2}. \quad (14)$$

ここで命題 1 より、 $\frac{d}{ds} Q_{v,\lambda,\beta}(s) = 0$ の解 s^* は $Q_{v,\lambda,\beta}(s)$ の最大値を与える。よって式 (14) において $\frac{d}{ds} Q_{v,\lambda,\beta}(s) = 0$ を解けば定理の結果が得られる。□

注意 2 図.6 で見られたように、 $Q_{v,\lambda,\beta}(s)$ の最大値を与える値 s^* は v に関して大きく変化しない。これにより、式 (13) で与えられた s^* を、一般の v に対して解析的に導出した結果であると近似的に定めることが可能である。この近似の評価結果については図 10 を用いて議論する。

注意 3 式 (13) で表された s^* は、 C, β のみに依存し、特に λ に依存していないことに注意する。特に、 s^* は C について比例の関係で増加するという性質は、図.7 で観察された結果と合致する。さらに、図.8 で観察された s^* の性質として、 s^* の λ の推移に対する値の変化はほとんど起こらないことが示されていたが、式 (13) で表された s^* は λ に依存しておらず、図.8 の結果と性質が合致することが分かる。

図.10 は、定理 2 により得られた s^* と数値計算により導出した s^* との比較結果である。なお、式 (9) で定められた条件をみたす範囲で v の値を推移させている。図.10 の結果より、特にパラメータ λ, β, C の値が小さい場合、 v が十分大きい場合であっても s^* の値の誤差は小さいままであることがわかる。一方で、 λ, β, C の値が大きい場合は、シミュレーション結果に対する誤差がある程度認められることに注意する。

5. 結論

本研究では、基地局の配置が homogeneous なポワソン点過程によって配置されるという Single-tier のセルラネットワークモデル上で、移動するユーザー端末がハンドオフを制限する時間をパラメータとして制御する Time-based なハンドオフ制限のモデルについて定義を与えた。さらに、定義された Time-based なハンドオフ制限モデルの下で、ユーザー端末のハンドオフレートと平均受信データレートの解析的な計算結果を与えた。そして、Time-based なハンドオフ制限モデルにおけるユーザー端末のスループット関数を、ハンドオフレートと平均受信データレートのトレードオフを表すよう定義し、パラメータ s に関するスループット関数の推移を観察した。その結果、一定条件下でスループット関数は s に関する最大値をもつことを数値計算により確認した。すなわち、一定のハンドオフ制限時間を設けることでユーザー端末のスループットが増加する可能性があることが、本研究によって明らかになった。さらに本研究では、スループット関数がハンドオフ制限時間 s に関して最大値を持つための必要十分条件を考察し、またこれらの最大値を与えるパラメータ s の値 s^* を解析的に計算するための手法についても提案した。

一方で、Time-based なハンドオフ制限モデルのハンドオフレートや受信データレートの計算では、それぞれ近似計算を導入したためにシミュレーション結果との誤差が確認された。本論文では、受信データレートの計算結果に見られる誤差の解消方法を提案したもの、ハンドオフレートにおける誤差の解消方法は未提案であり、今後の課題である。

参考文献

- [1] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri, and M. S. Alouini, “Handover management in dense cellular networks: A stochastic geometry approach,” IEEE ICC 2016, 10.1109/ICC.2016.7510709, 2016.
- [2] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri, and M. S. Alouini, “Velocity-aware handover management in two-tier cellular networks,” IEEE Trans. Wirel. Commun., 10.1109/TWC.2017.2655517, 2017.
- [3] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, 3rd ed. Wiley, 2013.
- [4] W. Bao and B. Liang, “Stochastic Geometric Analysis of Handoffs in User-centric Cooperative Wireless Networks,” IEEE INFOCOM 2016, 10.1109/INFOCOM.2016.7524592, 2016.
- [5] W. Bao and B. Liang, “Optimizing Cluster Size Through Handoff Analysis in User-Centric Cooperative Wireless Networks,” IEEE Trans. Wireless Commun., vol. 17, pp. 766–778, 2017.
- [6] K. Tokuyama, and N. Miyoshi, “Data rate and handoff rate analysis for user mobility in cellular networks,” IEEE WCNC 2018, 10.1109/WCNC.2018.8377167, 2018.
- [7] 徳山 喜一, “移動体通信におけるハンドオーバー制限のモデルとその解析,” 日本オペレーションズ・リサーチ学会 2018 年秋季研究発表会 アブストラクト集, pp84–85, 2018.
- [8] J. G. Andrews, F. Baccelli, and R. K. Ganti, “A tractable approach to coverage and rate in cellular networks,” IEEE Trans. Commun., vol. 59, pp. 3122–3134, 2011.

客のグルーピングのある待ち行列の安定条件: Foster の定理による導出

鶴見 星花[†] 松尾 容典^{*} 塩田 茂雄^{*}

[†] 千葉大学 工学部 都市環境システム学科

^{*} 千葉大学 大学院融合理工学府 都市環境システムコース

概要：本論文では，1 人ずつサービスを受ける客とグループを組んでサービスを受ける客の混在する待ち行列の安定条件を，フォスターの定理を用いて導出する．

1. はじめに

2 種類（タイプ 1，タイプ 2）の客が到着する待ち行列を考える．タイプ 1 とタイプ 2 の客は，それぞれ到着率 λ_1 および λ_2 のポアソン過程に従って到着する．タイプ 1 の客は 1 人ずつサービスを受け，タイプ 2 の客は最大 N 人のグループを組み，グループ全体でサービスを受ける．タイプ 1 の客は，到着時に行列の最後尾に並ぶ．タイプ 2 の客は，行列内にタイプ 2 の N 人未満のグループ（以下，未完成グループ）があればそのグループに入って並び，未完成グループがいなければ行列の最後尾に並ぶ．サービスは行列の先頭から順に行われ，タイプ 1 の客のサービス時間，およびタイプ 2 のグループ客のサービス時間はいずれも平均 $1/\mu$ の指数分布に従う．以前，我々は本モデルの安定条件をドリフト条件を用いて導出したが [1]，本稿では安定条件を Foster の定理 [2] を用いて導出する．以下では， $\rho_1 \stackrel{\text{def}}{=} \lambda_1/\mu$ および $\rho_2 \stackrel{\text{def}}{=} \lambda_2/\mu$ とする．また，タイプ 2 の未完成グループがサービス受けている最中に到着したタイプ 2 の客は未完成グループに合流できるものとする¹．

2. 客のグルーピングのある待ち行列の安定条件

2.1. Foster の定理

タイプ 2 の客の各グループを構成する最初の客（グループ先頭客）の到着時点に着目する．便宜上，時刻 0 以降に最初に到着したグループ先頭客を 0 番とし， i 番目のグループ先頭客が到着した時点での系内客数を X_i とすると， $\{X_i\}_{i=0}^{\infty}$ は既約な離散時間マルコフ連鎖となる．Foster の定理 [2] より，以下の結果が得られる．

定理 2.1. $L \geq 0$ および $\epsilon > 0$ が存在し，全ての $l \geq L$ について

$$E[X_1|X_0 = l] \leq l - \epsilon,$$

であれば， $\{X_i\}$ は正再帰的であり，したがって定常分布を持つ．

本稿では，「客のグルーピングのある待ち行列が安定である」ことを「マルコフ連鎖 $\{X_i\}$ が定常分布を有する」こととみなし，定理 2.1 の適用要件（ $L \geq 0$ および $\epsilon > 0$ が存在し，全ての $l \geq L$ について $E[X_1|X_0 = l] \leq l - \epsilon$ ）が成立するための条件を導く．

2.2. 主要結果

定理 2.2. $\rho_1 + \frac{\rho_2}{N} < 1$ であれば，客のグルーピングのある待ち行列は安定である．

証明. i 番目のグループ先頭客の到着時刻を T_i ， $(T_0, T_1]$ 間のタイプ 1 の到着客数を A_0^1 ， $(T_0, T_1]$ 間のタイプ 1 とタイプ 2 をあわせた退去客数を D_0 （タイプ 2 の客はグループ全体で 1 人とみなす）とする．次が成立する．

$$E[X_1|X_0 = l] - l = E[A_0^1|X_0 = l] - E[D_0|X_0 = l] + 1. \quad (1)$$

$(T_0, T_1]$ 間のタイプ 2 の到着客数を A_0^2 とすると， $A_0^2 \leq N$ および $E[A_0^1|X_0 = l, A_0^2 < N] \leq E[A_0^1|X_0 = l, A_0^2 = N]$ であることより

$$E[A_0^1|X_0 = l] = E[A_0^1|X_0 = l, A_0^2 = N]p(A_0^2 = N) + E[A_0^1|X_0 = l, A_0^2 < N]p(A_0^2 < N)$$

¹この条件は本質的でなく，合流できないものとしても結論は変わらない．

$$\begin{aligned}
&\leq E[A_0^1|X_0 = l, A_0^2 = N]p(A_0^2 = N) + E[A_0^1|X_0 = l, A_0^2 = N]p(A_0^2 < N) \\
&= E[A_0^1|X_0 = l, A_0^2 = N] = \frac{N\lambda_1}{\lambda_2}.
\end{aligned} \tag{2}$$

0 番目のグループ先頭客が時刻 T_1 においても系内に存在する事象を $\mathcal{A}$, その他の事象を $\mathcal{A}^c$ とすると,

$$E[D_0|X_0 = l] = E[D_0|X_0 = l, \mathcal{A}]p(\mathcal{A}) + E[D_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c) \geq E[D_0|X_0 = l, \mathcal{A}]p(\mathcal{A}). \tag{3}$$

(1) に (2) および (3) を代入して

$$E[X_1|X_0 = l] - l \leq \frac{N\lambda_1}{\lambda_2} - E[D_0|X_0 = l, \mathcal{A}]p(\mathcal{A}) + 1. \tag{4}$$

時刻 T_0 に到着したタイプ 2 の客から数えて, N 人目のタイプ 2 の客が到着した時刻を $\tilde{T}_N$ とすると,

$$E[\tilde{T}_N - T_0|X_0 = l] = E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}]p(\mathcal{A}) + E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c).$$

ここで $E[\tilde{T}_N - T_0|X_0 = l] = \frac{N}{\lambda_2}$ である (時刻 T_0 での系内客数とその後のタイプ 2 の客の到着は独立に定まる). また, 事象 $\mathcal{A}$ では時刻 T_0 以降 T_1 までに待ち行列は空にならず, したがって時刻 T_0 に生じた未完成グループは必ず完成することから, 事象 $\mathcal{A}$ においては $\tilde{T}_N = T_1$ である. 以上より

$$E[T_1 - T_0|X_0 = l, \mathcal{A}]p(\mathcal{A}) = \frac{N}{\lambda_2} - E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c). \tag{5}$$

$E[D_0|X_0 = l, \mathcal{A}] = \mu E[T_1 - T_0|X_0 = l, \mathcal{A}]$ であることに注意すると, (4) および (5) より次が得られる.

$$\begin{aligned}
E[X_1|X_0 = l] - l &\leq \frac{N\lambda_1}{\lambda_2} - \frac{N\mu}{\lambda_2} + 1 + \mu E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c) \\
&= -\frac{N\mu}{\lambda_2} \left(1 - \rho_1 - \frac{\rho_2}{N}\right) + \mu E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c).
\end{aligned} \tag{6}$$

$\mathcal{A}$ は系内にいる客が全て退去する前に, N 人以上のタイプ 2 の客が到着する事象に等しいことから

$$P(\mathcal{A}) = 1 - \int_0^\infty \sum_{k=0}^{N-1} \frac{(\lambda_2 t)^k}{k!} e^{-\lambda_2 t} \frac{(\mu t)^{l-1}}{(l-1)!} \mu e^{-\mu t} dt = 1 - \sum_{k=0}^{N-1} \binom{k+l-1}{k} p^k (1-p)^l.$$

ここで $p \stackrel{\text{def}}{=} \frac{\lambda_2}{\lambda_2 + \mu}$ である. 以上より

$$\lim_{l \rightarrow \infty} P(\mathcal{A}^c) = \lim_{l \rightarrow \infty} \sum_{k=0}^{N-1} \binom{k+l-1}{k} p^k (1-p)^l = 0.$$

したがって, 期待値の絶対連続性より, $E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c) = E[(\tilde{T}_N - T_0)\mathbf{1}(\mathcal{A}^c)|X_0 = l]$ は $l \rightarrow \infty$ の極限で 0 に収束する ($\mathbf{1}(\mathcal{A}^c)$ は $\mathcal{A}^c$ 内でのみ 1 を取る関数). 仮定 $(\rho_1 + \frac{\rho_2}{N} < 1)$ より, $\delta \stackrel{\text{def}}{=} \frac{N\mu}{\lambda_2} \left(1 - \rho_1 - \frac{\rho_2}{N}\right) > 0$ であるが, $E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c)$ が $l \rightarrow \infty$ の極限で 0 に収束することから, L を十分大きくとれば, 全ての $l \geq L$ に対して,

$$\mu E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c) \leq \frac{\delta}{2},$$

したがって

$$E[X_1|X_0 = l] - l \leq -\frac{N\mu}{\lambda_2} \left(1 - \rho_1 - \frac{\rho_2}{N}\right) + \mu E[\tilde{T}_N - T_0|X_0 = l, \mathcal{A}^c]p(\mathcal{A}^c) \leq -\delta + \frac{\delta}{2} = -\frac{\delta}{2}. \tag{7}$$

とできる. つまり定理 2.1 の適用要件が成立する. したがって, 待ち行列は安定である. $\square$

参考文献

- [1] 松尾容典, 鶴見星花, 塩田茂雄, “客のグルーピングのある待ち行列の安定条件,” 日本オペレーションズ・リサーチ学会 秋季研究発表会, 1-C-10, 2018.
- [2] 宮沢正清, 待ち行列の数理とその応用, 牧野書店, 2006.

短期貸出客を優先処理する図書館の貸出図書の待ち行列

白木 詩乃* 塩田 茂雄**
千葉大学工学部都市環境システム学科*
千葉大学大学院工学研究院**

概要：本研究では、貸出図書の予約待ちの現象を所蔵冊数を窓口数、予約件数を待ち行列長とする待ち行列でモデル化し、短期貸出客用の優先貸出図書を設けることの効果を評価した。その結果、短期貸出客用図書を設けることで、短期貸出客の待ち時間が大幅に減少すること、さらには一般客の待ち時間は若干増えるものの、全体の待ち時間も減少することが確認された。

1. はじめに

公共図書館における貸出予約件数は年々増加している。特にベストセラー本に対する予約が集中し、大きな図書館では1つのタイトルに数百件の予約がつくなど、これまでにない予約件数が数えられるようになった[1]。このように貸出予約件数が膨大となる問題に対して公共図書館が行っている対策として、同じタイトルの本を複数冊所蔵する複本が挙げられる。しかし新刊のベストセラーの本は数年経つと発売直後と比べ利用回数が減ってしまう。また図書館の資料購入費には限りがあるため、なるべく同じタイトルの本は購入したくないという意見もある。千葉市の公共図書館における貸出予約件数ベスト1の“蜜蜂と遠雷”という図書は2018年10月17日の時点で所蔵冊数49冊に対し、予約件数は1,040件となっていた[2]。千葉市の公共図書館の貸出期間は2週間が限度であるため、この状態から利用者が予約を入れると、1年ほど待つ可能性がある。

本研究では、貸出図書の予約待ちの現象を所蔵冊数を窓口数、予約件数を待ち行列長とする待ち行列でモデル化する。待ち行列理論では、サービス時間の短い客からサービスをする待ち時間が最小になることが知られている[3]。そこで、短期貸出客に優先的に貸し出す図書を設けることにより利用者全体の平均待ち時間を軽減する方策の有効性について検討する。短期貸出客用の優先図書のあるモデルは、複数の窓口を持ち、サービス時間分布の異なる2種類の客（優先客、非優先客）が到着し、優先客が優先的に使用できる窓口が設けられた待ち行列システムとみなすことができる。本研究では、シミュレーションによりこの待ち行列システムの性能を評価し、短期貸出客に優先的に貸し出す図書を設けることの有効性を定量的に調べることにする。

2. モデルとサービス規律

複数の窓口からなる待ち行列システムを考える。窓口の数を c とし、優先客が優先的に使用できる窓口（以下、優先客用窓口）の数を $s(\leq c)$ とする¹。優先客用以外の共通窓口は、優先客と非優先客の両方が公平に利用できるものとする。非優先客および優先客はそれぞれ到着率 λ_1 および λ_2 のポアソン過程に従って到着する。客は1つの窓口からサービスを受ける。各窓口でのサービス時間は互いに独立であり、非優先客は平均 $1/\mu_1$ をもつ指数分布に、優先客は平均 $1/\mu_2$ をもつ指数分布にそれぞれ従う($\mu_1 > \mu_2$)。

優先客が到着したとき、優先客用の窓口に空きがあればその窓口でサービスを受け、優先客用窓口に空きがない場合は共通窓口でサービスを受ける。非優先客が到着したときは、共通窓口に空きがあればその窓口でサービスを受ける。共通窓口に空きがない場合は、優先客用窓口に空きがある（したがって待機中の優先客がいない）ときに限り、そこでサービスを受ける。窓口に空きがない場合、優先客は優先客用行列に、非優先客は非優先客用行列にそれぞれ並ぶ。サービスが終了し客が退出して優先客用の窓口に空きができた場合、優先客用行列の先頭の客が空いた窓口案内される。このとき優先客用行列に客がいない場合は非優先客用行列の先頭の客が案内される。共通の窓口に空きができた場合、優先客用行列と非優先客用行列の先頭の客のうち、到着時刻の早い方が空いた窓口案内される。図1にシステムの概要を示す。

3. シミュレーション結果

窓口数を2018年10月17日時点の千葉市図書館における“蜜蜂と遠雷”の所蔵冊数（49）に合わせ、さらに $\lambda_1 = 4\lambda_2$ （非優先客は優先客の4倍）の条件のもとで、利用率が $\rho = 0.999$ となるように到着率 λ_1, λ_2 を設定してシミュレーションを行った。なお、非優先客の平均サービス時間は14日とし、優先客の平均サービス時

¹「優先客用窓口」は優先客が優先的に使用できるが、優先客が不在の時は非優先客も使用することができる。

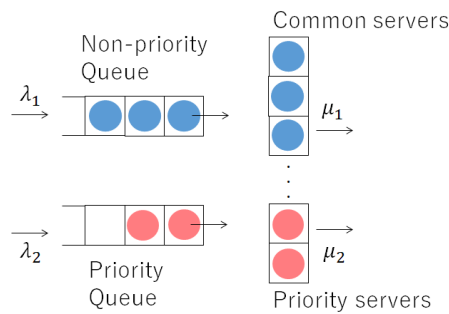
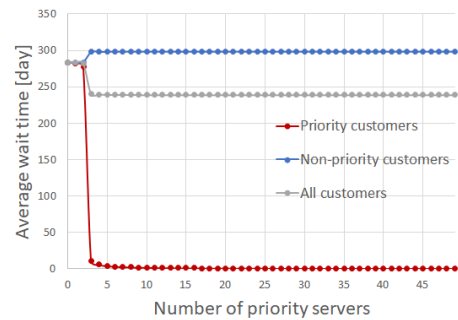
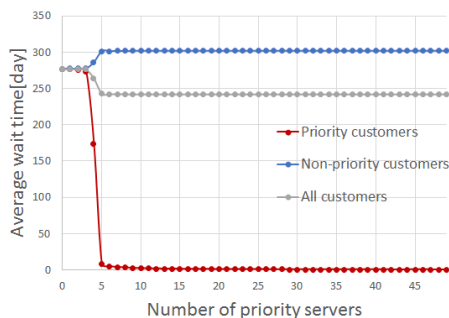


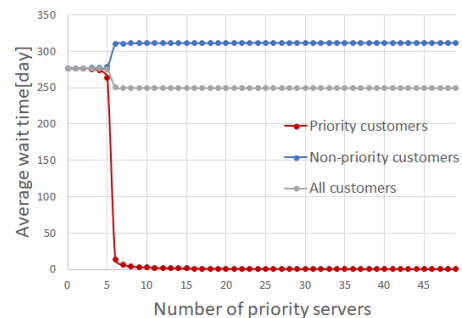
図 1: システムの概要



(a) 優先客の平均サービス時間 = 3 日



(b) 優先客の平均サービス時間 = 5 日



(c) 優先客の平均サービス時間 = 7 日

図 2: 優先客, 非優先客の平均待ち時間

間は 3 日, 5 日, 7 日の 3 通りの場合を想定した. この条件で, 優先客用の窓口数を変化させ, 優先客と非優先客の平均待ち時間を評価した. 結果を図 2(a)-(c) にそれぞれ示す. 図には全体の平均待ち時間も記載した.

優先客の平均サービス時間が 3 日の場合, 優先客用窓口数を 7 とすると, 優先客の平均待ち時間は約 99 % 減少し, 非優先客の平均待ち時間は約 5% 増加した. また, 全体の平均待ち時間は約 15% 減少した. 同様に, 優先客の平均サービス時間が 5 日 (7 日) の場合, 優先客用窓口数を 9 (10) とすると, 優先客の平均待ち時間は約 99 % 減少し, 非優先客の平均待ち時間は約 5% (8%) 増加し, 全体の平均待ち時間は約 15% (10%) 減少した. また, 優先客用窓口数をある程度大きくとると, 優先客, 非優先客の平均待ち時間は優先客用窓口数に対してほとんど依存しない結果となった.

4. まとめと今後の展望

本研究では, 貸出図書の予約待ちの現象を待ち行列モデルで表現し, 短期貸出客用の優先貸出図書を設けることの効果を評価した. その結果, 短期貸出客用の図書を設けることで, 短期貸出客の待ち時間は大幅に減少し, 一般客の待ち時間は若干増えるものの, 全体の待ち時間も減少することが確認された. この結果より, 短期貸出客用の図書を設けることは十分に効果があると結論づけられる.

本稿で取り上げたモデルは, 全てが共通窓口の場合は, $M/H_2/N$ モデルと等価になる. 全てが共通窓口の場合のシミュレーション結果が, $M/H_2/N$ モデルの結果と等しくなることを確認するとともに, 本稿で扱った優先客用窓口がある場合の待ち行列モデルを解析的に評価する手法についても今後検討したい.

参考文献

- [1] 宮部頼子, 現代図書館情報学シリーズ 4 図書館サービス概論, 樹村房, 2012.
- [2] 千葉市図書館, 予約ベスト一覧: 蔵書検索,
<https://www.library.city.chiba.jp/licsxp-opac/WOpacMsgRankMenuToMsgRankListAction.do>, 2018-10-17 閲覧.
- [3] 高橋幸雄, 森村英典, 経営科学のニューフロンティア 7 混雑と待ち, 朝倉書店, 2001.

飲食店における客席案内戦略に関する考察

野田 脩平 塩田 茂雄*

* 千葉大学 大学院融合理工学府 都市環境システムコース

概要：飲食店における客席の利用状態の変化を連続時間マルコフ連鎖でモデル化し、各種客席案内戦略を客収容効果の観点で評価した。席詰め案内がランダム案内と比べて、僅かな席占有数の増加がみられた。

1. はじめに

飲食店ではピーク時間帯になると、席を詰めて座るように促すことがある。しかし、実際に席を詰めることで客収容効果がどの程度改善するかは必ずしも明らかでない。本稿では、飲食店での客席案内を損失形待ち行列でモデル化し、客収容効果の観点から各種客席案内戦略の効率性を評価する。

2. 定式化

2.1. モデル

N 個のカウンタ席のある飲食店に n 人客 ($n = 1, 2, 3$) が到着率 λ_n のポアソン過程に従って来店する。 n 人客 ($n = 1, 2, 3$) の店内の滞在時間は平均 h_n の指数分布に従う。2 人客および 3 人客は、必ず隣り合ったカウンタ席に座るものとする。カウンタ席には左から順に 1 から N までの番号をふる。さらに、 n 人客が客席を占める各パターンに以下で定義する番号を付与する。

$$L_n \stackrel{\text{def}}{=} \sum_{k=1}^{N-n+1} 2^{k-1} S_k^{(n)}.$$

ここで $S_k^{(n)}$ は番号 k から $k+n-1$ までの座席を 1 組の n 人客が占める場合に 1、それ以外の場合は 0 をとる変数である¹。客席利用状態は 3 つの数字の組 (L_1, L_2, L_3) で識別できる。来店時に座れる席がなかった、もしくは案内を断られた客は、席が空くことを待たずに退去するものとする。

2.2. 案内戦略

客の来店時、もしくは退去時に客席利用状態 (L_1, L_2, L_3) の遷移が生じる。退去時の遷移先は一意に定まる。一方、客の来店時の遷移先には任意性があり、客の店内への案内方法に依存する。本稿では、この客の店内への案内方法を「客席案内戦略」と呼ぶこととし、特に以下の 4 つの客席案内戦略の効率性について考察する。

1. ランダム案内 案内可能な席の中から、ランダムに（等確率で）選んだ席に案内する。
2. 席詰め案内 案内可能な席の中で、最も左端に近い席に案内する。
3. 公平ランダム案内 3 人客の案内可能席が存在する場合にのみ、案内可能な席の中から、ランダムに選んだ席に案内する。
4. 公平席詰め案内 3 人客の案内可能席が存在する場合にのみ、案内可能な席の中で、最も左端に近い席に案内する。

ここで、3 と 4 は来店時に入店できない確率（ブロック率）が 1 人客、2 人客、3 人客の全てで等しくなる客席案内戦略である。客席案内戦略が定まると、客席利用状態 (L_1, L_2, L_3) 間の遷移が定まる。

2.3. 定常状態確率

時刻 t における客席利用状態 $(L_1(t), L_2(t), L_3(t))$ は有限状態連続時間マルコフ連鎖に従う。状態 (L_1, L_2, L_3) の定常状態確率を $p(L_1, L_2, L_3)$ で表し、状態を辞書式に並べた定常状態確率ベクトルを $\mathbf{p} = (p(0, 0, 0), p(1, 0, 0), \dots)$ により定める。客席案内戦略 s のもとでの推移率行列を $\mathbf{Q}_s$ とすると、客席案内戦略 s のもとでの定常状態確率 $\mathbf{p}_s$ は大域平衡方程式 $\mathbf{p}_s \mathbf{Q}_s = \mathbf{0}$ を満たす。大域平衡方程式を解くことにより $\mathbf{p}_s$ が求まり、 $\mathbf{p}_s$ からカウンタ席の平均利用数や n 人客のブロック率を算出できる。

¹例えば $S_1^{(2)}$ と $S_2^{(2)}$ が同時に 1 になることはないので、 L_2 は 0 から $2^{N-1} - 1$ までの全ての値を取るわけではない。同様に、 L_3 は 0 から $2^{N-2} - 1$ までの全ての値を取るわけではない。

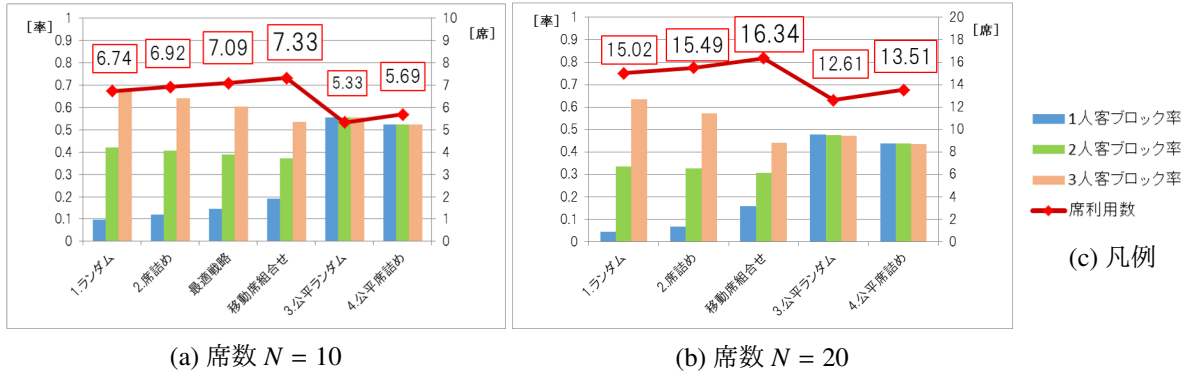


図 1: タイプ客別ブロック率と客席利用数

2.4. マルコフ決定過程による最適戦略の導出

本モデルは客席案内戦略に依存して推移率行列が定まることから、マルコフ決定過程により最適案内戦略を導出することが可能である [1]。例えば、目的関数をカウンタ席の平均利用数に取る。時刻 0 で客席利用状態 (L_1, L_2, L_3) から出発したときの時刻 t までの延べ席利用数の期待値 $G_{(L_1, L_2, L_3)}(t)$ は t が十分大きいとき、 $G_i(t) = gt + g_i$ となる。 $g_{(0,0,0)} = 0$ とおくと、 $\mathbf{g} \stackrel{\text{def}}{=} (g, g_{(1,0,0)}, g_{(2,0,0)}, \dots)^\top$ は $A\mathbf{g} = \mathbf{u}$ を満たす。行列 A の要素 a_{ij} 、ベクトル $\mathbf{u} \stackrel{\text{def}}{=} (u_{(0,0,0)}, u_{(1,0,0)}, u_{(2,0,0)}, \dots)^\top$ の要素は以下で定まる。

$$a_{ij} = \begin{cases} -q_{ij} & i \neq 0, \\ 1 & i = 0, \end{cases} \quad u_{(L_1, L_2, L_3)} = \sum_{n=1}^3 nh_n s_{(L_1, L_2, L_3)}(n)$$

ここで q_{ij} は推移率行列の要素であり、 $s_{(L_1, L_2, L_3)}(n)$ は客席利用状態 (L_1, L_2, L_3) において n 人客を入店させるときに 1、入店を断るときに 0 をとる変数である（案内戦略に依存する）。 $\mathbf{g}$ が求まれば、各客席利用状態 (L_1, L_2, L_3) で取るべき案内戦略が決まる。適当な初期案内戦略から出発して $\mathbf{g}$ を求め、 $\mathbf{g}$ に基づいて案内戦略を更新し、更新結果に基づいて $\mathbf{g}$ を求めることを繰り返すことにより（Howard の政策反復法 [2]）、最適客席案内戦略が求まる。

3. 数値実験

カウンタ 10 席と 20 席の両ケースについて、各客席案内戦略の平均客席利用数を評価した。なお、10 席の場合は数値計算で評価したが、20 席の場合は（状態数の爆発により数値計算が困難であったため）シミュレーションを評価に用いた。平均サービス時間 [時] は $(h_1, h_2, h_3) = (1/2, 3/4, 1)$ ，平均到着率 [組/時] は $(\lambda_1, \lambda_2, \lambda_3) = (6, 3, 1.5)$ （10 席），および $(\lambda_1, \lambda_2, \lambda_3) = (12, 6, 3)$ （20 席）とした。図 1 にカウンタ席の平均利用数（右軸）と客のブロック率（左軸）を示す。シミュレーション結果は時間 100 のシミュレーションの 100 回分の平均である。図 1 において「移動席組合せ」は店内の客の座席位置を変えて新しい客を案内できるスペースが生じ得るならば、常に客の入店を許可する方策をとった場合の結果であり、最大限客詰めして達し得る平均利用席数の上限値の目安として示した。なお、「移動席組合せ」の定常状態確率は積形式で与えられ、カウンタ 20 席の場合でも数値評価が可能である。

図より、席詰め案内、ランダム案内、また（10 席の場合）マルコフ決定過程による最適客席戦略による客席利用数の差は比較的小さいことが読み取れる。ただいずれの客席案内戦略においても、1 人客と 3 人客のブロック率には大きな差が生ずる。一方、ブロック率を公平にすると、席の平均利用数は大きく低下する。また、「移動席組合せ」方式の客席利用数は最大となるが、席詰め案内やランダム案内との差はさほど顕著ではなく、常識的な案内戦略で十分効果が得られることも確認された。

参考文献

- [1] 三宅 功, “異速度通信混在回線の最適回線留保制御,” 電子情報通信学会論文誌, vol.J71-B, no.12, pp.1419–1424, 1988.
- [2] A. Howard, Dynamic Programming and Markov Processes, M. I. T. Press, 1960.

M/G/1 待ち行列における待ち時間制約と滞在時間制約の比較

井上 文彰
大阪大学工学研究科
電気電子情報工学専攻

概要：本稿では、客の待ち時間に制約のある M/G/1 待ち行列 ($M/G/1+G^w$) ならびに客の滞在時間に制約のある M/G/1 待ち行列 ($M/G/1+G^s$) を考える。M/G/1+ G^w 待ち行列では、客は制約時間以内にサービスが開始しないとき、システムから即座に離脱する。また、客は一旦サービスを受け始めると、サービスの完了まではシステムから離脱しない。一方、M/G/1+ G^s 待ち行列では、客は制約時間以内にサービスが終了しないとき、サービスを受けている途中であっても、システムを即座に離脱する。本稿では、これら二種類の M/G/1 待ち行列の性能指標間に成り立つ関係を考察する。

1. はじめに

本稿では、客の途中退去が起こる二種類の待ち行列モデルを考える。一つ目のモデルでは、それぞれの客は待ち時間に制約をもち、経過待ち時間がその制約に達すると、客はサービスを受けずに即座にシステムから離脱する。このモデルでは、客は一旦サービスを受け始めると、サービスの完了まではシステムを離脱しない。一方、二つ目のモデルでは、それぞれの客は滞在時間に制約をもち、経過滞在時間がその制約に達すると、サービスの途中であっても即座にシステムから離脱する。

いずれのモデルについても、本稿では、客はポアソン過程に従って到着し、客のサービス時間は独立かつ同一に分布するものと仮定する。さらに、客の (待ち時間あるいは滞在時間に対する) 制約時間長は独立かつ同一に分布すると仮定する。[3] において導入された記法に倣って、本稿では、待ち時間制約のあるモデルを M/G/1+ G^w 待ち行列、滞在時間制約のあるモデルを M/G/1+ G^s 待ち行列と呼ぶことにする。また、M/G/1+ G^w 待ち行列ならびに M/G/1+ G^s 待ち行列をまとめて、M/G/1+ G 待ち行列と表記する。

本稿では、同一の到着率 λ ($\lambda > 0$)、同一のサービス時間分布、ならびに同一の制約時間長分布によって特徴づけられる M/G/1+ G^w 待ち行列ならびに M/G/1+ G^s 待ち行列を考察する。 H ならびに G をそれぞれ、サービス時間ならびに制約時間長を表す一般の確率変数とする。簡単のため、 $\Pr(H = 0) = \Pr(G = 0) = 0$ を仮定する。

便宜上、一貫して次の記法を用いることにする。任意の非負確率変数 Y に対し、 $Y(x) := \Pr(Y \leq x)$ ($x \geq 0$) をその確率分布関数とし、 $\bar{Y}(x) := 1 - Y(x)$ を補分布関数とする。また、 Y の確率密度関数が存在するとき、それを $y(x)$ ($x \geq 0$) と表す。さらに、M/G/1+ G^w 待ち行列に関する量には、上添字 'w' を付し、M/G/1+ G^s 待ち行列に関する量には上添字 's' を付す。

本稿では、時間制約の無い、すなわち、待ち時間や滞在時間の長さに関わらず、必ずサービスが完了するまでシステムを離脱しない客の存在を許すモデルを考える。そのような客は、無限大の制約時間長を有するため、制約時間長の分布は一般には劣確率的であるものとする。すなわち、 $g_\infty \in [0, 1)$ をランダムに選ばれた客が無限大の制約時間長を有する確率とすると

$$\lim_{x \rightarrow \infty} \bar{G}(x) = g_\infty$$

が成立する。一方で、サービス時間分布は有限の平均 $E[H] < \infty$ をもつことを仮定する。 $\rho := \lambda E[H]$ をトラヒック強度とし、

$$\rho g_\infty < 1 \tag{1}$$

が成り立つことを仮定する。[6, Theorem 2.1] より、(1) が成り立つとき、上記の M/G/1+ G^w 待ち行列ならびに M/G/1+ G^s 待ち行列はいずれも安定であることが保証される。

M/G/1+ G^w 待ち行列ならびに M/G/1+ G^s 待ち行列に対しては、数多くの解析結果が既に知られている [1, 3, 4, 8, 9, 10, 11]。特に、[9] では、M/G/1+ G^w 待ち行列ならびに M/G/1+ G^s 待ち行列の仮待ち時間分布はヴォルテラ第二種積分方程式の解として定められること、および、いずれのモデ

ルにおいてもロス確率 P_{loss} は仮待ち時間分布を用いて与えられることが示されている。 V を定常仮待ち時間を表す一般の確率変数とする。 まず、レベルクロッシング法 [5, 7] により、次式が得られる。

$$v^w(x) = \lambda \pi_0^w \bar{H}(x) + \lambda \int_0^x v^w(y) \bar{G}(y) \bar{H}(x-y) dy, \quad x > 0 \quad (2)$$

$$v^s(x) = \lambda \pi_0^s \bar{H}(x) \bar{G}(x) + \lambda \int_0^x v^s(y) \bar{H}(x-y) \bar{G}(x-y) dy, \quad x > 0 \quad (3)$$

ただし、 π_0 は定常状態においてシステムが空である確率を表し、次式を満たす。

$$\pi_0^w + \int_0^\infty v^w(x) dx = 1, \quad \pi_0^s + \int_0^\infty v^s(x) dx = 1 \quad (4)$$

(2) ならびに (3) は第二種ヴォルテラ積分方程式であり、これらと正規化条件 (4) から仮待ち時間分布は一意に定められる。さらに、PASTA より、それぞれのモデルにおけるロス確率 P_{loss} は

$$P_{\text{loss}}^w = \Pr(G \leq V^w) = \int_0^\infty v^w(x) G(x) dx \quad (5)$$

$$P_{\text{loss}}^s = \Pr(G \leq V^s + H) \quad (6)$$

$$= \int_{x=0}^\infty dG(x) \left(\pi_0^s \bar{H}(x) + \int_{y=0}^x v^s(y) \bar{H}(x-y) dy + \int_{y=x}^\infty v^s(y) dy \right) \quad (7)$$

で与えられる。(2) を用いると、(5) は次式で書き換えられる。

$$P_{\text{loss}}^w = 1 - \frac{1 - \pi_0^w}{\rho}. \quad (8)$$

すなわち、 $M/G/1+G^w$ 待ち行列におけるロス確率はシステムが空である確率 π_0^w を用いて与えられる。一方、 $M/G/1+G^s$ 待ち行列ではこのような単純な公式を得ることは困難であるように思われる。

本稿では、 $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列の性能指標間に成り立つ関係を考察する。大半の先行研究の主題は、 $M/G/1+G^w$ 待ち行列ならびに $M/G/1+G^s$ 待ち行列の個別のモデルに対する解析、あるいは、これら両方を包含するより一般的な制約メカニズムをもつ待ち行列 [4, 8, 9] の解析であり、 $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列の比較を行っている研究はほとんど存在しない。これら二種類のモデル間の比較に関する主要な文献は [17] である。この文献ではサービス時間が幾何分布に従う一般の離散時間待ち行列モデルにおいて $P_{\text{loss}}^w \leq P_{\text{loss}}^s$ が成り立つことが証明されている。その証明は、サービス時間が幾何分布に従うとき (8) と同様に P_{loss}^s が π_0^s を用いて与えられること [16]、ならびに $\pi_0^w \leq \pi_0^s$ がサンプルパスの観察により得られることに基づいている。これはサービス時間が幾何分布に従うという仮定が本質的となっているため、同様のアプローチによってこの結果を他のサービス時間分布に拡張することは容易ではない。なお、[15] に示されている通り、 $P_{\text{loss}}^w \leq P_{\text{loss}}^s$ は一般には必ずしも成立せず、この不等式は自明ではないことに注意を要する。

本稿では、最初に、 $M/G/1+G^s$ 待ち行列における定常仮待ち時間分布、定常待ち時間分布、ならびに定常滞在時間分布は、対応する $M/G/1+G^w$ 待ち行列における定常仮待ち時間分布、定常待ち時間分布、ならびに定常滞在時間分布を用いてそれぞれ与えられることを示す。次に、この結果を元に、 $M/G/1+G^w$ 待ち行列ならびに $M/G/1+G^s$ 待ち行列の性能指標間に成り立つ関係を考察する。特に、 $P_{\text{loss}}^w \leq P_{\text{loss}}^s$ が成り立つための十分条件を導出する。最後に、サービス時間が指数分布に従う場合 ($M/M/1+G$) と制約時間長が指数分布に従う場合 ($M/G/1+M$) を考え、これら特別なモデルでは、 P_{loss}^w と P_{loss}^s は単純な式によって関係づけられることを示す。

以下の本稿は次のように構成される。第 2 節では、 $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ を考察し、第 3 節ではそれらを特別化したモデルを考察する。最後に第 4 節において本稿を締めくくる。

2. $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列の関係

任意の確率変数 Y と事象 A に対し、 $[Y | A]$ で Y を A で条件付けて得られる確率変数を表すことにする。 W をサービスの開始までにシステムを途中退去しなかった客の実待ち時間を表す一般の確

率変数とする。また、 D をサービスの完了までシステムを途中退去しなかった客の滞在時間を表す一般の確率変数とする。PASTA より、次の関係式が成り立つことが確かめられる。

$$W^w =_{\text{st}} [V^w \mid V^w < G] \quad (9)$$

$$D^w =_{\text{st}} W^w + H \quad (10)$$

$$W^s =_{\text{st}} [V^s \mid V^s < G] \quad (11)$$

$$D^s =_{\text{st}} [V^s + H \mid V^s + H < G] \quad (12)$$

ただし、 $=_{\text{st}}$ は分布の意味での等価性を表す。

補題 1. 同一の到着率 λ ，同一のサービス時間分布 $H(x)$ ($x \geq 0$)，ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する M/G/1+G^w 待ち行列と M/G/1+G^s 待ち行列において、次式が成立する。

$$V^s =_{\text{st}} W^w \quad (13)$$

証明. (5) と (9) より、 W^w の分布は次式で特徴づけられる。

$$\Pr(W^w = 0) = \frac{\pi_0^w}{1 - P_{\text{loss}}^w}, \quad w^w(x) = \frac{v^w(x)\overline{G}(x)}{1 - P_{\text{loss}}^w}, \quad x > 0 \quad (14)$$

よって、(2) より $\Pr(W^w = 0)$ と $w^w(x)$ は次式を満たすことが確かめられる。

$$w^w(x) = \left(\lambda \Pr(W^w = 0) \overline{H}(x) + \lambda \int_0^x w^w(y) \overline{H}(x-y) dy \right) \overline{G}(x), \quad x > 0$$

上式は V^s の分布に関する積分方程式 (3) と同一である。 V^s の分布は (3) および正規化条件 (4) により一意に定められ、一方、 W^w の分布も同様に上式および正規化条件

$$\Pr(W^w = 0) + \int_0^\infty w^w(x) dx = 1$$

により一意に定められる。したがって、 $V^s =_{\text{st}} W^w$ が成立する。□

定理 1. 同一の到着率 λ ，同一のサービス時間分布 $H(x)$ ($x \geq 0$)，ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する M/G/1+G^w 待ち行列と M/G/1+G^s 待ち行列において、次式が成立する。

$$V^s =_{\text{st}} [V^w \mid V^w < G], \quad W^s =_{\text{st}} [W^w \mid W^w < G], \quad D^s =_{\text{st}} [D^w \mid D^w < G] \quad (15)$$

証明. (15) は (9), (10), (11), (12), および (13) から即座に導かれる。□

定理 1 は、定常な M/G/1+G^s 待ち行列における仮待ち時間分布、実待ち時間分布、ならびに滞在時間分布は、それと対応する定常な M/G/1+G^w 待ち行列における仮待ち時間分布、実待ち時間分布、ならびに滞在時間分布を用いてそれぞれ与えられることを示している。この結果の直接的な帰結として、確率順序を用いた次の関係式が得られる。

定義 1 ([13, Eq. (1.A.1)]). 非負確率変数 X ならびに Y に対し、

$$\Pr(X > x) \leq \Pr(Y > x) \quad \text{for all } x \geq 0$$

が成り立つとき、 X は通常の確率順序の意味で Y 以下であるといい、 $X \leq_{\text{st}} Y$ と表す。

注意 1. $X \leq_{\text{st}} Y$ が成り立つことは

$$E[\phi(X)] \leq E[\phi(Y)]$$

が (期待値の存在する) 任意の非減少関数 $\phi(x)$ について成り立つことの必要十分条件である [13, (1.A.7)]. 特に、 $X \leq_{\text{st}} Y \Rightarrow E[X] \leq E[Y]$ が成立する。

系 1. 同一の到着率 λ , 同一のサービス時間分布 $H(x)$ ($x \geq 0$), ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列において, 次式が成立する.

$$V^s \leq_{st} V^w, \quad W^s \leq_{st} W^w, \quad D^s \leq_{st} D^w$$

補題 1 の証明は付録 A に示す.

直観的には, 補題 1 より, $M/G/1+G^s$ は $M/G/1+G^w$ に比べて混雑の度合いが小さいということが示唆される. 特に, $V^s \leq_{st} V^w$ から $\pi_0^s \geq \pi_0^w$ が得られることがわかる. この結果から $P_{loss}^w \leq P_{loss}^s$ が成り立つことを予想するかもしれないが, 実際にはこれは一般には成り立たない (第 3 節で反例を示す). 以下では, $P_{loss}^w \leq P_{loss}^s$ が成り立つための十分条件を導出する.

補題 2. 同一の到着率 λ , 同一のサービス時間分布 $H(x)$ ($x \geq 0$), ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する定常な $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列を考える. これらのモデルにおけるロス確率は次式で与えられる.

$$P_{loss}^w = E[G(V^w)] \quad (16)$$

$$P_{loss}^s = E[G(D^w)] \quad (17)$$

証明. (16) は (5) から即座に導かれる. また, (17) は (6), (10), (13) から得られる. $\square$

注意 2. (17) より, $M/G/1+G^s$ 待ち行列におけるロス確率 P_{loss}^s は, 対応する $M/G/1+G^w$ 待ち行列においてサービスを受ける客の滞在時間が, それとは独立にサンプルされた制約時間長より長い確率と等しいことがわかる.

定義 2 ([12, Page 49]). 非負確率変数 X は

$$E[X - t \mid X \geq t] \leq E[X], \quad \forall t \geq 0$$

を満たすとき new better than used in expectation (NBUE) であるという.

補題 3. サービス時間分布が NBUE であるとき, $M/G/1+G^w$ 待ち行列において次式が成立する.

$$V^w \leq_{st} D^w \quad (18)$$

証明. $\tilde{H}$ をサービス時間の平衡分布に従う確率変数とする. すなわち, その密度関数は $\tilde{h}(x) = \bar{H}(x)/E[H]$ で与えられる. (8) ならびに (14) より, (2) の右辺は次式で書き換えられる.

$$v^w(x) = (1 - \pi_0) \left[\Pr(W^w = 0) \tilde{h}(x) + \int_0^x w^w(y) \tilde{h}(x - y) dy \right], \quad x > 0$$

これを確率変数間の関係として記述すると, 次式を得る.

$$[V^w \mid V^w > 0] =_{st} W^w + \tilde{H} \quad (19)$$

サービス時間分布 $H(x)$ が NBUE であるとき, 次式が成立する [13, Theorem 1.A.31].

$$\tilde{H} \leq_{st} H$$

よって, (10), (19), ならびに [13, Theorem 1.A.3] より,

$$V_+^w := [V^w \mid V^w > 0] =_{st} W^w + \tilde{H} \leq_{st} W^w + H =_{st} D^w \quad (20)$$

が成立する. したがって, 次のように補題 3 が導かれる.

$$\Pr(V^w > x) = (1 - \pi_0^w) \Pr(V_+^w > x) \leq \Pr(V_+^w > x) \leq \Pr(D^w > x), \quad x \geq 0$$

$\square$

定理 2. 同一の到着率 λ , 同一のサービス時間分布 $H(x)$ ($x \geq 0$), ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列を考える. サービス時間分布が NBUE であるとき, 次式が成立する.

$$P_{\text{loss}}^w \leq P_{\text{loss}}^s \quad (21)$$

証明. $G(x)$ は非減少関数であるため, (18) より,

$$E[G(V^w)] \leq E[G(D^w)]$$

を得る (注意 1 参照). したがって, (16) ならびに (17) より (21) を得る. □

3. 特別な場合

3.1. サービス時間が指数分布に従う場合 ($M/M/1+G$)

サービス時間がパラメタ μ ($\mu > 0$) の指数分布に従う場合を考える. この場合, 指数分布の無記憶性から $\tilde{H} =_{\text{st}} H$ を得る. したがって, (10) ならびに (19) より, 次式を得る ((20) を参照).

$$[V^w \mid V^w > 0] =_{\text{st}} D^w \quad (22)$$

定理 3. 同一の到着率 λ , 同一の平均サービス時間 $1/\mu$, ならびに同一の制約時間長分布 $G(x)$ ($x \geq 0$) を有する $M/M/1+G^w$ 待ち行列と $M/M/1+G^s$ 待ち行列において, 次式が成立する.

$$P_{\text{loss}}^s = \frac{P_{\text{loss}}^w}{\rho(1 - P_{\text{loss}}^w)} \quad (23)$$

証明. (16) および $G(0) = 0$ という仮定から,

$$P_{\text{loss}}^w = (1 - \pi_0^w)E[G(V^w) \mid V^w > 0]$$

が成立する. したがって, (8), (17), ならびに (22) より

$$P_{\text{loss}}^w = \rho(1 - P_{\text{loss}}^w)E[G(D^w)] = \rho(1 - P_{\text{loss}}^w) \cdot P_{\text{loss}}^s$$

が成立するため, これより (23) を得る. □

注意 3. 指数分布は NBUE であるため, 定理 2 より $P_{\text{loss}}^w \leq P_{\text{loss}}^s$ が成立する. (8) より $\rho(1 - P_{\text{loss}}^w) = 1 - \pi_0^w < 1$ であることに注意すると, この不等式は (23) から直接確かめられる.

注意 4. $M/M/1+G^w$ 待ち行列において, π_0^w は次式で与えられる [10, 14].

$$\pi_0^w = \left(1 + \lambda \int_{x=0}^{\infty} \exp \left[-\mu x + \lambda \int_{y=0}^x \bar{G}(y) dy \right] dx \right)^{-1}$$

よって (8) から P_{loss}^w が求められ, さらに, (23) から P_{loss}^s を求めることができる.

3.2. 制約時間長が指数分布に従う場合 ($M/G/1+M$)

制約時間長がパラメタ γ ($\gamma > 0$) の指数分布に従う場合を考える. 任意の非負確率変数 Y に対し, $\mathcal{L}[Y](s) := E[e^{-sY}]$ ($\text{Re}(s) > 0$) を, そのラプラス・スティルチェス変換 (LST) とする. このとき, (16) ならびに (17) より次式を得る.

$$P_{\text{loss}}^w = 1 - \mathcal{L}[V^w](\gamma) \quad (24)$$

$$P_{\text{loss}}^s = 1 - \mathcal{L}[D^w](\gamma) \quad (25)$$

定理 4. 同一の到着率 λ , 同一のサービス時間分布 $H(x)$ ($x \geq 0$), ならびに同一の平均制約時間長 $1/\gamma$ を有する $M/G/1+G^w$ 待ち行列と $M/G/1+G^s$ 待ち行列において, 次式が成立する.

$$P_{\text{loss}}^s = 1 - \frac{\mathcal{L}[H](\gamma)}{\mathcal{L}[\tilde{H}](\gamma)} \left(1 - \frac{P_{\text{loss}}^w}{\rho(1 - P_{\text{loss}}^w)} \right) \quad (26)$$

証明. (19) より, 次式が成り立つことが確かめられる.

$$\mathcal{L}[V^w](s) - \pi_0^w = (1 - \pi_0^w) \cdot \mathcal{L}[W^w](s) \cdot \mathcal{L}[\tilde{H}](s)$$

上式において $s = \gamma$ とおくと, (24) より

$$1 - P_{\text{loss}}^w - \pi_0^w = (1 - \pi_0^w) \cdot \mathcal{L}[W^w](\gamma) \cdot \mathcal{L}[\tilde{H}](\gamma)$$

が得られ, さらに, この式と (8) から次式が得られる.

$$\mathcal{L}[W^w](\gamma) = \frac{1}{\mathcal{L}[\tilde{H}](\gamma)} \left(1 - \frac{P_{\text{loss}}^w}{\rho(1 - P_{\text{loss}}^w)} \right)$$

したがって, (10) より

$$\mathcal{L}[D^w](\gamma) = \mathcal{L}[W^w](\gamma) \cdot \mathcal{L}[H](\gamma)$$

であることに注意すると, (25) から (26) が導かれる. $\square$

注意 5. サービス時間が指数分布に従うとき $\mathcal{L}[\tilde{H}] = \mathcal{L}[H]$ が成立するため, (26) は (23) に単純化される.

注意 6. M/G/1+M^w 待ち行列では, π_0^w は次式で与えられる [8, 10].

$$\pi_0^w = \left(1 + \sum_{n=1}^{\infty} \rho^n \prod_{k=1}^{n-1} \mathcal{L}[\tilde{H}](k\gamma) \right)^{-1}$$

よって, (8) から P_{loss}^w が求められ, さらに, (26) から P_{loss}^s を求めることができる (注意 4 も参照).

サービス時間が平衡超指数分布に従う M/H₂/1+M^w 待ち行列ならびに M/H₂/1+M^s 待ち行列におけるロス確率を, サービス時間の変動係数 $\text{Cv}[H]$ の関数としてプロットした結果を図 1 に示す. 図より, P_{loss}^w は $\text{Cv}[H]$ の増加に伴って単調に増加しており, これはサービス時間分布の変動が凸順序 [13, Eq. (3.A.1)] に関して増加するのに伴って M/G/1+G^w 待ち行列のロス確率が増加するという理論的結果 [10] と整合している. 一方, P_{loss}^s は $\text{Cv}[H]$ がある程度大きくなると減少に転じ, 結果としてサービス時間の変動が比較的大きい領域では $P_{\text{loss}}^w > P_{\text{loss}}^s$ が成立している. すなわち, これは $P_{\text{loss}}^w \leq P_{\text{loss}}^s$ に対する反例を与えている.

4. まとめ

本稿では M/G/1+G^w 待ち行列と M/G/1+G^s 待ち行列の性能指標間の関係を論じた. 最初に, M/G/1+G^s 待ち行列における定常仮待ち時間分布, 定常待ち時間分布, ならびに定常滞在時間分布は, 対応する M/G/1+G^w 待ち行列における定常仮待ち時間分布, 定常待ち時間分布, ならびに定常滞在時間分布を用いてそれぞれ与えられることを示した. 次に, この結果を用いて, これらのモデルの性能指標間に成り立つ関係式を導出した. 最後に, サービス時間が指数分布に従う場合 (M/M/1+G) と制約時間長が指数分布に従う場合 (M/G/1+M) を考え, このとき, M/G/1+G^w 待ち行列と M/G/1+G^s 待ち行列のロス確率は単純な等式を通じて直接的に関係づけられることを示した.

A. 補題 1 の証明

補題 1 を示すには, $\Pr(X \leq Z) > 0$ を満たす任意の独立な確率変数 X ならびに Z に対し次式が成り立つことを示せば十分である.

$$X \leq_{\text{st}} [X \mid X \leq Z] \quad (27)$$

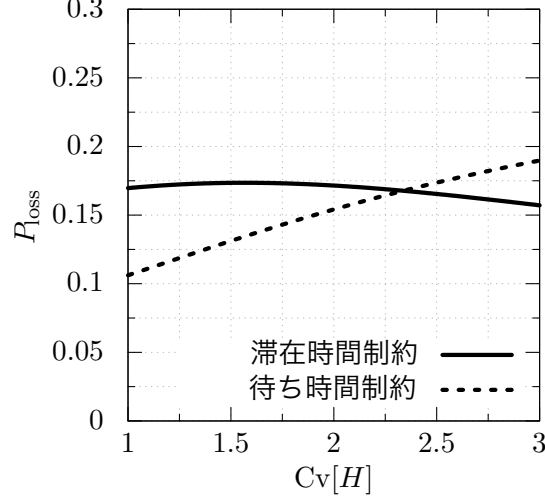


図 1: M/H₂/1+M 待ち行列における P_{loss}^w と P_{loss}^s ($\lambda = 0.7$, $E[H] = 1.0$, $E[G] = 10$).

定義より, $x \geq 0$ に対し次式を得る.

$$\begin{aligned}
& \Pr(X > x) - \Pr(X > x \mid X \leq Z) \\
&= \frac{\Pr(X \leq Z) \Pr(X > x) - \Pr(X > x, X \leq Z)}{\Pr(X \leq Z)} \\
&= \frac{\Pr(X > x, X > Z) - \Pr(X > Z) \Pr(X > x)}{\Pr(X \leq Z)} \\
&= \frac{1}{\Pr(X \leq Z)} \int_0^\infty \text{Cov}[\mathbb{1}_{\{X > x\}}, \mathbb{1}_{\{X > y\}}] dZ(y)
\end{aligned}$$

ただし, $\mathbb{1}_{\{\cdot\}}$ は指示関数を表し, $\text{Cov}[\cdot, \cdot]$ は共分散を表す. $\mathbb{1}_{\{w > t\}}$ ($t \geq 0$) は w の非減少関数であるため, 共分散不等式 [2] より

$$\text{Cov}[\mathbb{1}_{\{X > x\}}, \mathbb{1}_{\{X > y\}}] \geq 0, \quad x \geq 0, y \geq 0$$

が成立する. したがって, 上式より (27) が得られる.

謝辞

本研究の一部は JSPS 科研費 JP18K18007 の支援を受けている.

参考文献

- [1] Baccelli F., Boyer, P., and Hebuterne, G. (1984). Single-server queues with impatient customers. *Adv. Appl. Prob.* **16**, 887–905.
- [2] Behboodian, J. (1994). Covariance inequality and its applications. *Int. J. Math. Educ. Sci. Technol.* **25**, 643–647.
- [3] Boxma, O., Perry, D., and Stadge, W. (2011). The M/G/1+G queue revisited. *Queueing Systems* **67**, 207–220.
- [4] Brandt, A. and Brandt, M. (2013). Workload and busy period for M/G/1 with a general impatience mechanism. *Queueing Systems* **75**, 189–209.
- [5] Brill, P. H. and Posner, M. J. M. (1977). Level crossings in point processes applied to queues: Single-server case. *Operat. Res.* **25**, 662–674.

- [6] Browne, S. and Sigman, K. (1992). Work-modulated queues with applications to storage processes. *J. Appl. Prob.* **29**, 699–712.
- [7] Cohen, J. W. (1977). On up- and downcrossings. *J. Appl. Prob.* **14**, 405–410.
- [8] Daley, D. J. (1965). General customer impatience in the queue GI/G/1. *J. Appl. Prob.* **2**, 186–205.
- [9] Kovalenko, I. N. (1961). Some queueing problems with restrictions. *Theory Prob. Appl.* **6**, 205–208.
- [10] Inoue, Y. and Takine, T. (2015). Analysis of the loss probability in the M/G/1+G queue, *Queueing Systems* **80**, 363–386.
- [11] Inoue, Y. and Takine, T. (2015). The M/D/1+D queue has the minimum loss probability among M/G/1+G queues. *Operat. Res. Lett.* **43**, 629–632.
- [12] Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*, John Wiley & Sons, Chichester, UK.
- [13] Shaked, M. and Shanthikumar, J. G. (2007), *Stochastic Orders*, Springer, New York, NY.
- [14] Stanford, R. E. (1979). Reneging phenomena in single channel queues. *Math. Operat. Res.* **4**, 162–178.
- [15] van Houdt, B., Lenin, R. B., and Blondia, C. (2003). Delay distribution of (im)patient customers in a discrete time D-MAP/PH/1 queue with age-dependent service times. *Queueing Systems* **45**, 59–73.
- [16] van Velthoven, J., van Houdt, B., and Blondia, C. (2005). Response time distribution in a D-MAP/PH/1 queue with general impatience. *Stoch. Models* **21**, 745–765.
- [17] van Velthoven, J., van Houdt, B., and Blondia, C. (2006). On the probability of abandonment in queues with limited sojourn and waiting times. *Operat. Res. Lett.* **34**, 333–338.

メッセージ分割が発生する時のペイロード長が平均応答時間に及ぼす影響 ーバースト的にパケットが到着する場合ー

池川隆司
早稲田大学・東京大学

概要：本論文では、メッセージを複数のパケットに分割し、分割パケットが通信回線にバースト的に到着する時、ペイロード長が平均応答時間に及ぼす影響を考察する。このために、メッセージ長の分布とペイロード長が与えられた時、通信回線への分割パケットの集団到着とパケット長列の相関を表現している $M^X/G/1$ キューを提案する。長さが対数正規分布に従う Web オブジェクトを転送する数値例より、低負荷時、パケット当たりの平均応答時間の曲線はペイロード長に対して凹状となることを示す。さらに、メッセージ長が指数分布に従いつつ負荷がない場合、平均応答時間を最小にするペイロード長の近似値は、 $\ell^{(m)}$ を平均メッセージ長、 $\ell^{(h)}$ をヘッダ/トレーラー長とすると、 $\sqrt{2\ell^{(m)}\ell^{(h)}/3}$ で与えられることを示す。

1. はじめに

インターネットに代表されるパケット転送網では、ペイロード長（転送可能な最大パケット長）、ウィンドウサイズのような、QoS (Quality of Service) 尺度（平均応答時間、グッドプット等）に影響を与える様々な制御パラメータが存在する（[1, Section 1.4], [2, Section 1.2]）。制御パラメータの QoS 尺度への最適化は、通信網設計・管理において、重要な課題となっている。例えば、一般的なパケット転送網 [3]、無線網 [4]、工場用無線センサー網 [5]、無線網制御システム [6] において、この最適化問題を取り扱っている。

通信回線でのパケット（回線上で転送されるデータ単位）の転送時間は、パケット長を通信回線容量（速度）で割ることによって得られる [1, pp. 64–65]。さらに、パケットがビット誤りとなる確率はパケット長にほぼ比例する [3, p. 132]。したがって、パケット長の最大値を制限するペイロード長は、重要な制御パラメータの 1 つとなる。

メッセージ（Web ページ転送のようなアプリケーションによって発生するデータ単位）の長さは、ペイロード長より大きくなることが多い。このようなメッセージを通信網を使って転送させるために、送信局はメッセージ分割機能を備えている。例えば、通信プロトコル TCP/IP [7] や IEEE 802.11 MAC (media access control) では、メッセージ分割機能を規定している。送信局は、メッセージ分割機能によって、ペイロード長より長いメッセージを複数のデータに分割する。さらに、各データにヘッダ/トレーラーを付加し、パケットを作成する。

パケットがビット誤りにより廃棄された時、廃棄パケットが誤り回復機能により再送される通信網においては、次に示すトレードオフが発生するため、グッドプット（受信局によって単位時間当たりに正しく受信される平均ビット数）を最適化するペイロード長が存在する [8–11]。

- ペイロード長が非常に小さい場合、ビット誤りによるパケット廃棄はほとんど発生しないが、パケット当たりのオーバーヘッド（ヘッダ/トレーラー）が無視できないため、グッドプットは小さくなってしまう。
- ペイロード長が非常に大きい場合、パケットは頻繁にビット誤りとなる。その結果、多数の無駄なパケット再送を引き起こしてしまい、グッドプットは小さくなってしまう。

よって、パケットがビット誤りにより廃棄され、廃棄パケットが再送されるシナリオでは、グッドプットを最大化するペイロード長が存在する。

メッセージ分割が発生するシナリオでは、QoS 尺度をパケットの平均応答時間とすると、他のトレードオフが存在する。ここで、平均応答時間は、バッファでの平均待ち時間と平均転送時間に分解することができる。

- ペイロード長が非常に大きい場合、メッセージ分割はほとんど発生しない。そのため、パケット到着のバースト性により生じる平均待ち時間は十分小さい。一方、パケットの平均転送時間は、メッセージのそれとほぼ等しいため、メッセージ分割が発生した場合と比べて大きくなる。
- メッセージ分割の発生率を十分小さくするペイロード長より、ペイロード長が小さくなると、メッセージ当たりの平均分割パケット数は増加していく。その結果、パケット到着のバースト性が増し、パケットの平均待ち時間は増大する。一方、メッセージ分割により平均パケット長は小さくなるため、平均転送時間は小さくなっていく。結果的に、平均待ち時間は増加していくが、相対的に平均転送時間は小さくなるため、平均応答時間は減少していく。
- ペイロード長が非常に小さい場合、メッセージ当たりの平均分割パケット数は著しく増加し、平均待ち時間は平均転送時間と比べて極めて大きくなる。その結果、平均応答時間は大きくなってしまう。

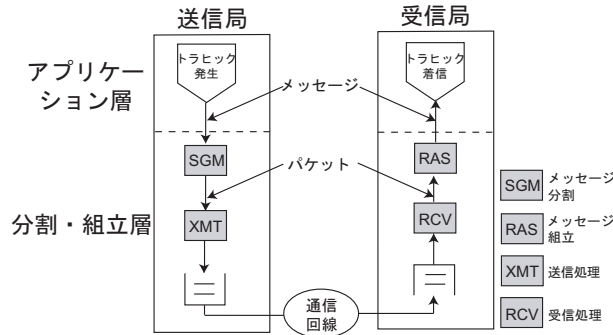


図 1: 通信網モデル

上記の議論より、メッセージ分割が発生する場合、平均応答時間とペイロード長の関係には、トレードオフが発生すると予想される。

今まで後者のトレードオフについては、深く考察されていない。筆者の知る限り、メッセージ長分布とペイロード長が陽に与えられ、メッセージ分割が発生する状況のもとでのトレードオフの分析は行われていない。本論文の目的は、このトレードオフを、集団到着待ち行列システムを使って、詳細に分析することである。

集団到着待ち行列システムは、分割パケット（待ち行列システムでの客に相当する）の集団的な到着の振舞を表現することができる。しかし、通常取り扱われている集団到着待ち行列システム [12], [13, Section 9.3] では、メッセージが複数のパケットに分割された場合、パケットの転送時間（待ち行列システムでのサービス時間に相当する）には相関が生じるが、パケットの転送時間は互いに独立であると仮定している。本論文では、分割されたパケット長列の性質をもとに、この仮定を緩めている。

本論文の構成は次の通りである。次節において、本論文で取り扱う通信網のモデルを説明する。3 節では、分割パケット長の列を解析する。4 節では、3 節で得られた解析結果をもとに、パケット長とメッセージ当たりの分割パケット数の定常分布を導出する。5 節では、4 節で求めた定常分布の統計パラメータ（平均、分散等）を用いて、集団到着待ち行列システムの代表例である $M^X/G/1$ キューを使って通信網を表現し、その $M^X/G/1$ キューを解くことによって、平均応答時間の解を導出する。6 節では、メッセージ長が対数正規分布と指数分布に従う時、ペイロード長が平均応答時間に及ぼす影響を考察する。最後に、7 節では、本論文をまとめるとともに、今後の課題に言及する。

2. 通信網モデル

本節では、最初に本論文で取り扱う 2 階層の通信網モデルを説明する。次に解析の簡単化のために導入した仮定を述べる。

2.1. 2 階層モデル

図 1 に示すように、各端末はアプリケーション層と分割・組立層の 2 つの層から構成される。

アプリケーション層: メッセージと呼ぶデータ単位を発生させるトラヒックの発生源を含む。

分割・組立層: メッセージ分割・機能を実装する。メッセージ分割機能は、ペイロード長 $\ell^{(d)} (> 0)$ を超えるメッセージを、複数のデータに分割させる。そのデータを情報フィールドにカプセル化し、ヘッダ/トレーラーのようなプロトコル制御情報 (PCI: Protocol Control Information) を付け加えたデータ単位 (パケットと呼ぶ) を生成する。さらに、通信回線を使ってパケットを転送させる。

2.2. 仮定

解析の簡単化のため、メッセージ長と PCI 長に関して、以下の仮定を設ける。

A1: $L_i^{(m)} (> 0)$ を送信局のアプリケーション層で発生した i 番目のメッセージ長とする。確率変数列 $\{L_i^{(m)}; i = 1, 2, \dots\}$ は、同一分布 $F^{(m)}(\cdot)$ に従い、互いに独立 (i.i.d.: independent and identically distributed) である。

A2: メッセージ長の分布 $F^{(m)}(\cdot)$ は、有限の平均値 $\ell^{(m)}$ と分散 $\{\sigma^{(m)}\}^2$ を持つ。

A3: PCI の長さは一定長 $\ell^{(h)}$ である。

3. パケット長列モデル

メッセージ分割機能により 1 つのメッセージから生成された (複数の) パケットの列を図 2 に示す。図 2 に示すように、 $L_i^{(m)} > \ell^{(d)}$ の時、 $i (= 1, 2, \dots)$ 番目メッセージは複数のパケットに分割される。 X_i を i 番目メッ

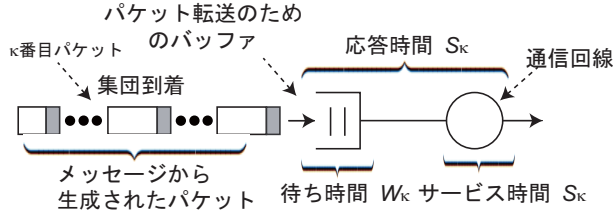


図 4: 待ち行列システム

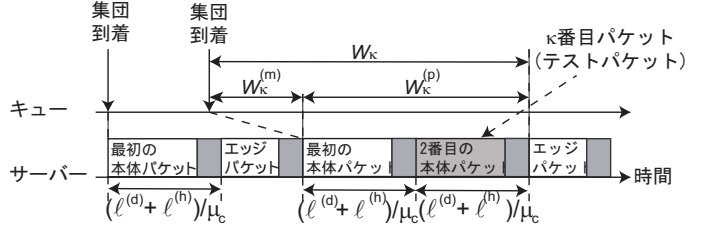


図 5: $W_k^{(m)}$ と $W_k^{(p)}$ の関係

エッジパケット長の定常分布 $F^{(E)}(\cdot)$ は、

$$F^{(E)}(x) = \begin{cases} 0, & 0 \leq x < \ell^{(h)}, \\ \sum_{n=0}^{\infty} \left(F^{(m)}(x + n\ell^{(d)} - \ell^{(h)}) - F^{(m)}(n\ell^{(d)} - \ell^{(h)}) \right), & \ell^{(h)} \leq x \leq \ell^{(d)} + \ell^{(h)}, \\ 1, & x > \ell^{(d)} + \ell^{(h)}, \end{cases} \quad (5)$$

として与えられる。

$\ell^{(p)}$ を $\{L_k^{(p)}\}$ の定常分布の平均値とする。仮定 A1 と A2 より、 $\ell^{(p)}$ は有限となる。(2)–(5) 式より、次式を得る。

$$\ell^{(p)} = \int_0^{\infty} x dF^{(p)}(x) = \pi^{(E)} \ell^{(m)}. \quad (6)$$

なお、(4)–(6) 式の導出については、文献 [14] を参照して欲しい。

4.2. メッセージ当たりの分割パケット数の定常分布

分割パケット数列 $\{X_i\}$ の定常分布を $\{\beta_n (\triangleq \Pr.(X_i = n))\} (n = 1, 2, \dots)$ で表す。 β_n は、

$$\beta_n = \Pr. \left((n-1)\ell^{(d)} < L^{(m)} \leq n\ell^{(d)} \right) = F^{(m)}(n\ell^{(d)}) - F^{(m)}((n-1)\ell^{(d)}) = u_{n-1} - u_n, \quad (7)$$

として与えられる。

$\bar{\beta}$ と $\bar{\beta}^{(2)}$ を、それぞれ定常分布 $\{\beta_n\}$ を持つ確率変数 X_i の平均と 2 次モーメントとする。 $\bar{\beta}$ と $\bar{\beta}^{(2)}$ は、(3) 式と (4) 式より、

$$\bar{\beta} = \sum_{n=0}^{\infty} n \beta_n = \sum_{n=0}^{\infty} u_n = \frac{1}{\pi^{(E)}}, \quad (8)$$

$$\bar{\beta}^{(2)} = \sum_{n=0}^{\infty} n^2 \beta_n = \sum_{n=0}^{\infty} (2n+1) u_n, \quad (9)$$

として与えられる。ここで、仮定 A1 より、 $\{X_i\}$ は i.i.d. の確率変数の集合となる。

5. 平均応答時間

確率変数 R_k を k 番目パケットの応答時間とする。確率変数列 $\{R_k\}$ の定常状態での平均 $E[R_k]$ の解を導出するために、図 4 に示す単一サーバー待ち行列システムを使って、通信網を表現する。

図 4 で示した待ち行列システムでのサーバーは、通信回線に対応する。 k 番目客（つまり k 番目パケット）のサービス時間 S_k は、パケットの転送時間と等しい。つまり、 μ_c を通信回線の容量（速度）とすると、 S_k は、

$$S_k = \frac{L_k^{(p)} + \ell^{(h)}}{\mu_c}, \quad (10)$$

として与えられる。

応答時間 R_k は、待ち行列システムでの用語を用いると、 k 番目パケットがキューへ到着してからサービスを完了するまでの時間として定義される。応答時間 R_k は、待ち時間（キューで過ごす時間） W_k とサービス時間 S_k の和となる。

分割パケットのバースト的な到着を表現するために、1 つのメッセージから生成される分割パケットの集団が同時に通信回線に到着する（つまり集団到着）を仮定する。

解析の簡単化のために、以下の仮定を導入する。

B1: 待ち行列システムへの集団は、到着率 λ のポアソン過程に従って到着する。

B2: 待ち行列システム内で収容可能な最大客数は、無限である。

B3: 客は、FIFO (First-In, First-Out) 規則に従って送信される。

B4: 負荷 (トラフィック密度) ρ は $\rho < 1$ を満たす³。

仮定 B1 – B3 より、待ち行列システムを $M^X/G/1$ キューとして取り扱うことができる。ここで、 X は集団の大きさを意味する。

仮定 B4 より、 $M^X/G/1$ キューは安定となる。

$M^X/G/1$ キューは多くの研究者により解析されている (例えば、[12], [13, Section 9.3])。本論文では、[12], [13, Section 9.3] 等によって頻繁に解析されてきた $M^X/G/1$ キューを、「通常 $M^X/GI/1$ キュー」と呼ぶことにする⁴。

通常 $M^X/GI/1$ キューでは、以下の仮定を設けている。

C1: 客 (パケットに相当) のサービス時間の列は、共通の確率分布に従い、互いに独立である。本論文では、これを i.i.d. サービス時間仮定と呼ぶ。

C2: 同一集団内の客はランダム規律に従いサービスされる。

実際のシステムでは、仮定 C1 と C2 は成立しない。例えば、分割されたパケット長の列は、互いに独立でない ((1) 式参照)。つまり、同一集団内の客のサービス時間の列は、i.i.d. ではなくなる。さらに、集団内の客における実際のサービス順序は、ランダム規律ではなく FIFO 規律である (仮定 B3 参照)。

これらの問題点を解決するために、相関がある分割パケット長列と FIFO 規律に従うパケット送信を表現する $M^X/G/1$ キューを提案する。

通常 $M^X/GI/1$ キューと同様に、 κ 番目パケットにおいて、確率変数 W_κ を独立な 2 つの確率変数 $W_\kappa^{(m)}$ と $W_\kappa^{(p)}$ に分解する (図 5 参照)。

$W_\kappa^{(m)}$: 集団の待ち時間。集団の到着からその集団の最初のパケット (客) の送信が開始されるまでの時間を意味する。

$W_\kappa^{(p)}$: κ 番目パケット (テストパケット) の待ち時間。テストパケットを有する集団内の最初のパケットの送信開始からそのテストパケットの送信開始までの時間を意味する。

本論文で対象とする通信網を表現した $M^X/G/1$ キューを解くことによって、次式を得る。

$$E[R_\kappa] = E[S_\kappa] + E[W_\kappa], \quad (11)$$

$$E[S_\kappa] = \frac{\ell^{(p)} + \ell^{(h)}}{\mu_c}, \quad (12)$$

$$E[W_\kappa] = E[W_\kappa^{(m)}] + E[W_\kappa^{(p)}], \quad (13)$$

$$E[W_\kappa^{(m)}] = \frac{\lambda}{2(1-\rho)} \frac{[\{\sigma^{(m)}\}^2 + \{\ell^{(m)}\}^2 + \bar{\beta} (2\ell^{(m)}\ell^{(h)} + \{\sigma^{(h)}\}^2) + \bar{\beta}^{(2)} \{\ell^{(h)}\}^2]}{\mu_c^2}, \quad (14)$$

$$\rho = \frac{\lambda (\ell^{(m)} + \bar{\beta} \ell^{(h)})}{\mu_c}, \quad (15)$$

$$E[W_\kappa^{(p)}] = \frac{(\ell^{(d)} + \ell^{(h)}) (\bar{\beta} - 1)}{2\mu_c}. \quad (16)$$

(11) – (16) 式の導出については、付録 A を参照して欲しい。

例 1 指数分布の場合

次式で示すように、メッセージ長が平均 $\ell^{(m)}$ の指数分布に従う場合を考える。

$$F^{(m)}(x) = 1 - e^{-\frac{x}{\ell^{(m)}}}, \quad x > 0. \quad (17)$$

u を $e^{-\frac{\ell^{(d)}}{\ell^{(m)}}}$ とすると、(3), (17) 式より、次式を得る。

$$u_n = e^{-\frac{n\ell^{(d)}}{\ell^{(m)}}} = u^n, \quad n = 0, 1, 2, \dots \quad (18)$$

³1 つのサーバーのキューの場合、 ρ はサーバーの利用率となる。

⁴仮定 C1 で述べるように、サービス時間は互いに独立であると仮定しているため、サービス時間の特徴を表す記号では “G” の代わりに “GI” を用いる。

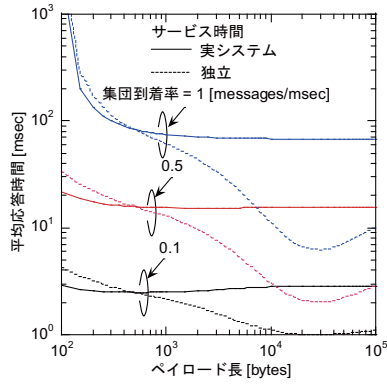


図 6: メッセージ長が対数正規分布に従う場合、集団到着率 λ が 0.1, 0.5, 1.0 [messages/msec] の時のペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$

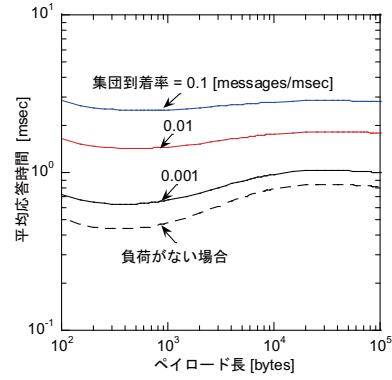


図 7: メッセージ長が対数正規分布に従う場合、集団到着率 λ が 0, 0.001, 0.01, 0.1 [messages/msec] の時のペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$

この時、(12), (14), (15), (16) 式内の項 $\ell^{(p)}, \sigma^{(m)}, \bar{\beta}, \bar{\beta}^{(2)}$ は、

$$\ell^{(p)} = (1 - u) \ell^{(m)}, \quad \sigma^{(m)} = \ell^{(m)}, \quad \bar{\beta} = \frac{1}{1 - u}, \quad \bar{\beta}^{(2)} = \frac{1 + u}{(1 - u)^2}, \quad (19)$$

として与えられる。

6. 数値例と考察

本節では、5 節で述べた結果を用いて、ペイロード長 $\ell^{(d)}$ が平均応答時間 $E[R_\kappa]$ に及ぼす影響を調べる。最初に、数値例での仮定を説明する。次に、長さが近似的に対数正規分布に従う Web オブジェクト (メッセージに対応) を転送するシナリオで、 $\ell^{(d)}$ が $E[R_\kappa]$ に及ぼす影響を考察する。最後に、 $E[R_\kappa]$ の陽解が得られる、メッセージ長が指数分布に従う場合のシナリオを想定する。

6.1. 仮定

IEEE 802.11g LAN (つまり通信回線の容量 $\mu_c = 54$ Mbps) を使って、パケットを転送するシナリオを考える。PCI の長さ $\ell^{(h)}$ は 38 バイトと仮定する。

6.2. メッセージ長が対数正規分布に従う場合

Web オブジェクトの長さは通常、ロングテールの性質を持つ。本論文では、測定結果 [15] に従い、次に示す対数正規分布を使って、Web オブジェクトの長さを近似する。

$$F^{(m)}(x) = \int_{y=0}^{y=x} \frac{1}{\sqrt{2\pi}\sigma_w y} e^{-\frac{(\log y - \mu_w)^2}{2\sigma_w^2}} dy, \quad x > 0. \quad (20)$$

ここで、メッセージ長の平均 $\ell^{(m)}$ と標準偏差 $\sigma^{(m)}$ は、それぞれ 4827 と 41,008 バイトとなるため [15]、パラメータ μ_w と σ_w は、それぞれ 6.34 と 2.07 となる。

様々な集団到着率 λ において、ペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$ を図 6 に示す。この例では、 λ として 0.1, 0.5, 1.0 [messages/msec] を用いた。この図では、i.i.d. サービス時間の仮定により導かれる誤差を評価するために、通常 $M^X/GI/1$ キューを使って得られる平均応答時間 $E[\hat{R}_\kappa]$ もプロットしている (注 1 参照)。 $E[\hat{R}_\kappa]$ の導出については、付録 B を参照して欲しい。

図 6 から以下の特徴を有することがわかる。

- D1: ペイロード長 $\ell^{(d)}$ を無限に大きくすると、 $E[R_\kappa]$ は有限値に近づく、
- D2: ペイロード長 $\ell^{(d)}$ が小さくなるほど、 $E[R_\kappa]$ は大きくなる、
- D3: 低負荷の時、 $E[R_\kappa]$ はペイロード長 $\ell^{(d)}$ に対して凹状となる。

特徴 D1 と D2 を深く考察するために、ペイロード長 $\ell^{(d)}$ が非常に大きい場合と小さい場合を考える。

- ペイロード長が非常に大きい時

この場合、メッセージ分割がほとんど発生しないため、以下を得る。

- $n = 1, 2, \dots$ の時の u_n がほぼ 0 となり、(4) 式より、エッジパケット発生率 $\pi^{(E)}$ はほぼ 1 となる。よって、各パケットの情報フィールド長はメッセージ長とほぼ同一となる。

- 集団サイズの集合 $\{X_i\}$ において、定常分布 $\{\beta_1, \beta_2, \beta_3, \dots\}$ は、 $\{1, 0, 0, \dots\}$ で近似できる。その結果、 $\{X_i\}$ の平均 $\bar{\beta}$ と 2 次モーメント $\bar{\beta}^{(2)}$ は、ほとんど 1 となる。よって、バースト的なパケット到着は発生しない。

以上より、ペイロード長 $\ell^{(d)}$ が非常に大きい時の待ち行列システムは、サービス時間をメッセージと PCI からなるデータ単位（この場合、パケットと同一視）の転送時間とする M/G/1 キューとほぼ等価となる。

M/G/1 キューの解析より、ペイロード長 $\ell^{(d)}$ が十分大きい時の平均応答時間 $\lim_{\ell^{(d)} \rightarrow \infty} E[R_\kappa]$ は、

$$\lim_{\ell^{(d)} \rightarrow \infty} E[R_\kappa] = E[S_\kappa] + \frac{\lambda E[S_\kappa^2]}{2(1 - \lambda E[S_\kappa])}, \quad (21)$$

として与えられる。ここで、 $E[S_\kappa]$ と $E[S_\kappa^2]$ は、

$$E[S_\kappa] = \frac{\ell^{(m)} + \ell^{(h)}}{\mu_c}, \quad E[S_\kappa^2] = \frac{\{\sigma^{(m)}\}^2 + (\ell^{(m)} + \ell^{(h)})^2}{\mu_c^2}, \quad (22)$$

である。

仮定 A2 より $E[S_\kappa]$ と $E[S_\kappa^2]$ は有限となるため、(21) 式と (22) 式より、 $\lim_{\ell^{(d)} \rightarrow \infty} E[R_\kappa]$ は有限となる。

- ペイロード長が非常に小さい場合

任意の n において $\lim_{\ell^{(d)} \rightarrow 0} u_n = \lim_{\ell^{(d)} \rightarrow 0} \int_n^\infty dF^{(m)} = 1$ となるため、次式を得る。

$$\lim_{\ell^{(d)} \rightarrow 0} \bar{\beta} = \lim_{\ell^{(d)} \rightarrow 0} \sum_{n=0}^\infty u_n = \sum_{n=0}^\infty \lim_{\ell^{(d)} \rightarrow 0} u_n \rightarrow \infty. \quad (23)$$

上式は、メッセージ分割が頻繁に発生することから容易に正当化できる。よって、

$$\lim_{\ell^{(d)} \rightarrow 0} W_\kappa^{(m)} \rightarrow \infty, \quad \lim_{\ell^{(d)} \rightarrow 0} W_\kappa^{(p)} \rightarrow \infty, \quad (24)$$

が得られる。上式は、ペイロード長 $\ell^{(d)}$ が非常に小さい時、パケットの平均待ち時間 $E[W_\kappa]$ が非常に大きいことを意味する。よって、平均応答時間 $E[R_\kappa]$ が非常に大きくなる。

特徴 D3 を明らかにするため、低負荷時でペイロード長 $\ell^{(d)}$ が平均応答時間 $E[R_\kappa]$ に及ぼす影響を精査する。負荷がない時（つまり、集団到着率 λ が 0）に加え、集団到着率 λ が 0.001, 0.01, 0.1 [messages/msec] の時のペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$ を図 7 に示す。図 7 より、平均応答時間 $E[R_\kappa]$ は最小値を持つことがわかる。

この理由を明らかにするために、負荷がない時の（仮想的な）平均応答時間 $E[R_\kappa](E[R_\kappa]_{\lambda=0})$ と表記）を考える。 $E[R_\kappa]_{\lambda=0}$ は、 $E[W_\kappa^{(m)}] = 0$ に注意すると、

$$E[R_\kappa]_{\lambda=0} = E[S_\kappa] + E[W_\kappa^{(p)}] \quad (25)$$

$$= \frac{\pi^{(E)} \ell^{(m)} + \ell^{(h)}}{\mu_c} + \left(\frac{\ell^{(d)} + \ell^{(h)}}{2\mu_c} \right) \left(\frac{1 - \pi^{(E)}}{\pi^{(E)}} \right), \quad (26)$$

として与えられる。

(26) 式から得られる以下の理由より、 $E[R_\kappa]_{\lambda=0}$ の曲線は凹状となる。

- $\pi^{(E)}$ はペイロード長 $\ell^{(d)}$ に対して単調増加である⁵、
- (26) 式の右辺の最初の項 ($E[S_\kappa]$ に対応する項) は、 $\pi^{(E)}$ に比例する、
- (26) 式の右辺の 2 番目の項 ($E[W_\kappa^{(p)}]$ に対応する項) は、 $\pi^{(E)}$ に対して凹状の性質を持つ。

注 1 通常 $M^X/GI/1$ キューを使った時の近似誤差

図 6 より、通常 $M^X/GI/1$ キューを用いた解析では、大きなペイロード長の時、多大な誤差をもたらすことがわかる。これは、通常 $M^X/GI/1$ キューが、サービス時間が一定でなく、かつ独立でない振舞を正確に表現できないことによる。一方、ペイロード長が小さい時、サービス時間が独立である場合と非独立である場合の平均応答時間の差分は小さくなる。これは、パケット長がほとんど固定値となる（つまり実際に送信されたパケットの集合では本体パケットが支配的となる）ことに起因している。

⁵ メッセージ長が指数分布に従う場合は、任意の $\ell^{(d)} (> 0)$ において $\frac{\partial \pi^{(E)}}{\partial \ell^{(d)}} = \frac{e^{-\frac{\ell^{(d)}}{\ell^{(m)}}}}{\ell^{(m)}} > 0$ となるため、容易に確かめられる。

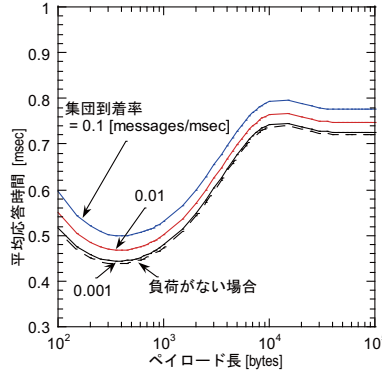


図 8: メッセージ長が指数分布に従う時、様々な集団到着率 λ におけるペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$

6.3. メッセージ長が指数分布に従う場合

最小の平均応答時間となるペイロード長を、 $\underline{\ell}^{(d)} (\triangleq \arg \min_{\ell^{(d)}} E[\hat{R}_\kappa])$ と表記する。 $\underline{\ell}^{(d)}$ の解を求めるために、メッセージ長が指数分布に従う場合を考える。

メッセージ長が指数分布に従う時、集団到着率 λ が 0, 0.001, 0.01, 0.1 [messages/msec] の場合のペイロード長 $\ell^{(d)}$ に対する平均応答時間 $E[R_\kappa]$ を図 8 に示す。この図より、メッセージ長が指数分布に従う時の平均応答時間 $E[R_\kappa]$ の傾向は、メッセージ長が対数正規分布に従う場合と同様であることがわかる。

(19) 式を使うことによって、 $E[R_\kappa]$ の陽解を求めることができる。しかし、その解は複雑である。そこで、 $E[R_\kappa]$ の陽解が簡易となる、負荷がない場合（つまり $\lambda = 0$ ）を考える。この場合の $E[\hat{R}_\kappa]$ を $E[\hat{R}_\kappa]_{\lambda=0}$ と表記すると、 $E[\hat{R}_\kappa]_{\lambda=0}$ は、

$$E[\hat{R}_\kappa]_{\lambda=0} = \frac{(1-u)\ell^{(m)} + \ell^{(h)}}{\mu_c} + \frac{u(\ell^{(d)} + \ell^{(h)})}{2(1-u)\mu_c}, \quad (27)$$

として与えられる。

$\lambda = 0$ の時、平均応答時間を最小とするペイロード長を $\underline{\ell}_{\lambda=0}^{(d)}$ と表記する。 $\underline{\ell}_{\lambda=0}^{(d)}$ は、(27) 式を $\ell^{(d)}$ で微分することにより得られ、 $\frac{\partial E[\hat{R}_\kappa]_{\lambda=0}}{\partial \ell^{(d)}}$ は、

$$\frac{\partial E[\hat{R}_\kappa]_{\lambda=0}}{\partial \ell^{(d)}} = 1 + \frac{\ell^{(m)} \left(1 - e^{-\frac{\ell^{(d)}}{\ell^{(m)}}} \right) - \ell^{(d)} - \ell^{(h)}}{2\ell^{(m)} \left(1 - e^{-\frac{\ell^{(d)}}{\ell^{(m)}}} \right)^2}, \quad (28)$$

として与えられる。

$E[\hat{R}_\kappa]_{\lambda=0}$ の必要条件は、 $\frac{\partial E[\hat{R}_\kappa]_{\lambda=0}}{\partial \ell^{(d)}} = 0$ である。(28) 式の右辺は非線形である。しかし、数値解析を使って $\underline{\ell}_{\lambda=0}^{(d)}$ の値を求めることができる。この例では、 $\underline{\ell}_{\lambda=0}^{(d)}$ は 366 バイトである。

$e^{-\frac{\ell^{(d)}}{\ell^{(m)}}} \approx 1 - \frac{\ell^{(d)}}{\ell^{(m)}} + \frac{1}{2} \left(\frac{\ell^{(d)}}{\ell^{(m)}} \right)^2$ を (28) 式に代入すると、 $\underline{\ell}_{\lambda=0}^{(d)}$ の近似式は、

$$\underline{\ell}_{\lambda=0}^{(d)} \approx \sqrt{\frac{2\ell^{(m)}\ell^{(h)}}{3}}, \quad (29)$$

として与えられ、この例での近似値は 350 バイトとなる。

7. 結論

本論文では、メッセージ分割が発生する通信網において、ペイロード長がパケットの平均応答時間に及ぼす影響を考察した。これを行うために、分割パケットのバースト的な到着とメッセージ当たりのパケット転送時間列の相関を表現する $M^X/G/1$ キューを提案した。この $M^X/G/1$ キューを解くことによって、メッセージ長の分布、ペイロード長、メッセージの到着率が与えられた時、平均応答時間を導出できることを示した。数値例より、平均応答時間は低負荷時、ペイロード長に対して凹状となることを示した。特に、メッセージ長が平均 $\ell^{(m)}$ の指数分布に従い、ヘッダ/トレーラ長を $\ell^{(h)}$ とすると、低負荷時、平均応答時間を最小にするペイロード長は近似的に $\sqrt{2\ell^{(m)}\ell^{(h)}/3}$ として与えられることを示した。

今後の課題は、複数のトラヒック源、通信回線にビット誤りが発生する場合への拡張や、ネットワーク条件が動的に変化する時の簡易で効果的なペイロード長最適化アルゴリズムの検討があげられる。

付録

A. (11) – (16) 式の導出

本付録では、(11) – (16) 式の導出を説明する。書籍 [12], [13, Section 9.3] で説明されているように、M^X/G/1 キューは、M/G/1 キューの解析を拡張することで容易に解析できる。

- (11) – (13) 式の導出

R_κ , S_κ , W_κ の定義から、(11) – (13) 式は明らかである。

- (14) – (16) 式の導出

- 準備

(14) – (16) 式を導出するために、確率変数 $L_n^{(\text{tp})}$ を導入する。 $L_n^{(\text{tp})}$ は、1 つのメッセージから n 個のパケットが生成される時、 n 個分のパケットの全長を意味する。

1 つのメッセージから生成されるパケットの情報フィールド長の総和 $\sum_{j=1}^n L_{ij}^{(\text{p})}$ は $L_i^{(\text{m})}$ と等しいことに注意すると、 $L_n^{(\text{tp})}$ は、

$$L_n^{(\text{tp})} = \sum_{j=1}^n L_{ij}^{(\text{p})} + n \ell^{(\text{h})} = L_i^{(\text{m})} + n \ell^{(\text{h})}, \quad (30)$$

として与えられる。よって、 $E[L_n^{(\text{tp})}]$ と $E[\{L_n^{(\text{tp})}\}^2]$ は、

$$E[L_n^{(\text{tp})}] = \ell^{(\text{m})} + n \ell^{(\text{h})}, \quad E[\{L_n^{(\text{tp})}\}^2] = \{\sigma^{(\text{m})}\}^2 + \{\ell^{(\text{m})}\}^2 + 2n \ell^{(\text{m})} \ell^{(\text{h})} + n^2 \{\ell^{(\text{h})}\}^2, \quad (31)$$

として与えられる。

- (14) 式

M/G/1 キューの議論 (例えば [2, Section 4.3.3]) より、 $E[W_\kappa^{(\text{m})}]$ は、

$$E[W_\kappa^{(\text{m})}] = \frac{\lambda E[\{S^{(\text{b})}\}^2]}{2(1-\rho)}, \quad (32)$$

として与えられる。ここで、 $S^{(\text{b})}$ は、1 つの集団 (つまり 1 つのメッセージから生成されるパケットの集まり) において、すべての客をサービスするために必要な総時間である。

(32) 式内の $E[\{S^{(\text{b})}\}^2]$ は、

$$E[\{S^{(\text{b})}\}^2] = \sum_{n=1}^{\infty} \text{Pr.}(X_i = n) E[\{S^{(\text{b})}\}^2 | X_i = n] = \frac{\sum_{n=1}^{\infty} \beta_n E[\{L_n^{(\text{tp})}\}^2]}{\mu_c^2}, \quad (33)$$

として与えられる。(31), (32), (33) 式から、(14) 式を得る。

- (15) 式

負荷 ρ の定義から、

$$\rho = \lambda E[S^{(\text{b})}], \quad (34)$$

を得る。 $E[S^{(\text{b})}]$ は、

$$E[S^{(\text{b})}] = \sum_{n=1}^{\infty} \text{Pr.}(X_i = n) E[S^{(\text{b})} | X_i = n] = \frac{\sum_{n=1}^{\infty} \beta_n E[L_n^{(\text{tp})}]}{\mu_c}, \quad (35)$$

として与えられる。(35) 式を (34) 式に代入すると、(15) 式を得る。

- (16) 式

* 1 つのメッセージが n 個のパケットに分割される時、パケットの集まりは連続する $(n-1)$ 個の本体パケットと 1 つのエッジパケットから構成される (図 2 参照)、

* 本体パケットの長さは $\ell^{(\text{d})} + \ell^{(\text{h})}$ である、

* 1つのメッセージから作られるパケットの集まり（集団）での r 番目に位置するテストパケットは、 $(r-1)$ 個の本体パケットの送信が完了するまで待たなければならない、
に注意すると、

$$E \left[W_{\kappa}^{(p)} \right] = \sum_{n=1}^{\infty} \Pr. (X_i = n) E \left[W_{\kappa}^{(p)} \mid X_i = n \right] = \sum_{n=1}^{\infty} \beta_n \frac{(n-1) (\ell^{(d)} + \ell^{(h)})}{2 \mu_c}, \quad (36)$$

となり、上式は (16) 式と一致する。

B. 通常 $M^X/GI/1$ キューから得られる平均応答時間の導出

平均応答時間 $E[\hat{R}]$ の導出において、i.i.d. サービス時間仮定の影響を受ける項は、負荷と集団の平均待ち時間である。それらを、それぞれ $\tilde{\rho}$ と $E[\tilde{W}_{\kappa}^{(m)}]$ で表す。

付録 A より、 $\tilde{\rho}$ と $E[\tilde{W}_{\kappa}^{(m)}]$ はそれぞれ $E[S^{(b)}]$ と $E[\{S^{(b)}\}^2]$ に依存することがわかる。i.i.d. サービス時間の仮定のもとで、 $E[S^{(b)}]$ と $E[\{S^{(b)}\}^2]$ を求めるために、1つのメッセージから n 個のパケットが発生し、パケット長が互いに独立である時のパケット長の総和を意味する確率変数 $\tilde{L}_n^{(tp)}$ を導入する。i.i.d. サービス時間の仮定より、 $E[\tilde{L}_n^{(tp)}]$ と $E[\{\tilde{L}_n^{(tp)}\}^2]$ は、次式で与えられる。

$$E \left[\tilde{L}_n^{(tp)} \right] = n \left(\ell^{(p)} + \ell^{(h)} \right), \quad E \left[\{\tilde{L}_n^{(tp)}\}^2 \right] = n^2 \left(\ell^{(p)} + \ell^{(h)} \right)^2 + n \left(\{\sigma^{(p)}\}^2 + \{\sigma^{(h)}\}^2 \right). \quad (37)$$

ここで、上式の $\{\sigma^{(p)}\}^2$ はパケット長の分散である。その陽解は、[14, Proposition 8] を参照して欲しい。付録 A と同様な手順を使うと、平均応答時間 $E[\hat{R}]$ の解を得る。

謝辞

本研究の一部は、JSPS 科研費 JP15K00139 の支援を受けました。(29) 式の導出において、東京大学大学院数理科学研究科中安淳氏から貴重なコメントを頂きました。記して感謝の意を表します。

参考文献

- [1] J. Kurose and K. Ross: “Computer Networking: A Top-Down Approach”, Pearson Education Kimited, seventh edition (2016).
- [2] H. Kobayashi and B. Mark: “System Modeling and Analysis: Foundations of System Performance Evaluation”, Prentice Hall (2008).
- [3] M. Schwartz: “Telecommunication Networks: Protocols, Modeling and Analysis”, Addison-Wesley Publishing Company (1987).
- [4] B. Fu, Y. Xiao, H. Deng and H. Zeng: “A survey of cross-layer designs in wireless networks”, IEEE Communications Surveys & Tutorials, **16**, 1, pp. 110 – 126 (2014).
- [5] V. C. Gungor and G. P. Hancke: “Industrial wireless sensor networks: Challenges, design principles, and technical approaches”, IEEE Transactions on Industrial Electronics, **56**, 10, pp. 4258–4265 (2009).
- [6] P. Park, S. C. Ergen, C. Fischione, C. Lu and K. H. Johansson: “Wireless network design for control systems: A survey”, IEEE Communications Surveys & Tutorials, **20**, 2, pp. 978–1013 (2018).
- [7] W. R. Stevens: “TCP/IP Illustrated, Volume 1: The Protocols”, Addison-Wesley Publishing Company (1994).
- [8] E. Modiano: “An adaptive algorithm for optimizing the packet size used in wireless ARQ protocols”, Journal Wireless Networks, **5**, 4, pp. 279–286 (1999).
- [9] G. Ciccicarese, M. D. Blasi, P. Marra, C. Palazzo and L. Patrono: “An algorithm for controlling packet size in IEEE 802.16e networks”, Computer Networks, **55**, 13, pp. 2873 – 2885 (2011).
- [10] W. Dong, C. Chen, X. Liu, Y. He, Y. Liu, J. Bu and X. Xu: “Dynamic packet length control in wireless sensor networks”, IEEE Transactions on Wireless Communications, **13**, 3, pp. 1172–1181 (2014).
- [11] D. Kim, H. Nam and D. Kim: “Adaptive code dissemination based on link quality in wireless sensor networks”, IEEE Internet of Things Journal, **4**, 3, pp. 685–695 (2017).
- [12] P. J. Burke: “Delays in single-server queues with batch input”, Operations Research, **23**, 4, pp. 830–833 (1975).
- [13] H. C. Tijms: “A First Course in Stochastic Models”, John Wiley & Sons, Ltd. (2003).
- [14] T. Ikegawa, Y. Kishi and Y. Takahashi: “Data-unit-size distribution model when message segmentations occur”, Performance Evaluation, **69**, 1, pp. 1–16 (2012).
- [15] M. Molina, P. Castelli and G. Foddis: “Web traffic modeling exploiting TCP connections’ temporal clustering through HTML-REDUCE”, IEEE Network, **14**, 3, pp. 46–55 (2000).