複数連鎖待ち行列ネットワークに対するマルコフ連鎖モンテカルロ法
——クライアント・サーバモデルの性能評価

数理第五研究室 杉村 由花; 指導教員 杉原 厚吉 教授，松井 知己 教授

2007 年 7 月 30 日

1 はじめに

本研究では，複数連鎖待ち行列に対してマルコフ連鎖モンテカルロ法 (MCMC 法) を適用し，クライアント・サーバモデルの性能評価を行う．

本研究では 2 種類のモデル化を行う．一つめのモデルは，各連鎖の客数が 1 のモデルである．このモデルにおいて，客数 m に対し $\mathrm{O}(m^2)$ の計算時間でシステムの正規化定数が厳密に計算できることを示した．また，システムの定常分布からのサンプリングを CFTP アルゴリズムにより行うため，システムの定常分布と等しい定常分布を持ち，単調性というよい性質を持つマルコフ連鎖 $\mathcal{M}_{\mathrm{S}}$ を人工的に構成した．$\mathcal{M}_{\mathrm{S}}$ を用いた CFTP アルゴリズムの性能を調べるため，計算機実験を行った．

二つめのモデルは，各連鎖の客数が一般のモデルであり，パラメータの似通ったジョブを同じ連鎖に属させる．一つめのモデルにおいて，計算時間は客数 n に依存するが，その一部をパラメータの種類数 m への依存に置き換えることで計算時間の短縮を狙うものである．正規化定数の計算を行うため，サーバに滞在する客数 h を固定することによる定常分布の条件付き分布を考え，それを定常分布に持つ人工的マルコフ連鎖 $\mathcal{M}^h_{\mathrm{M}}$ を構成した．また，$\mathcal{M}^h_{\mathrm{M}}$ の収束が速いことを示した．よって，$\mathcal{M}^h_{\mathrm{M}}$ を用いて MCMC 法を行うことにより，$\mathcal{M}^h_{\mathrm{M}}$ の正規化定数を推定でき，いくつかの h について MCMC 法を行うことで残りの h についても正規化定数を推定できれば，システム全体の正規化定数を精度保証つきで推定できると期待される．

2 準備

2.1 クライアント・サーバモデル

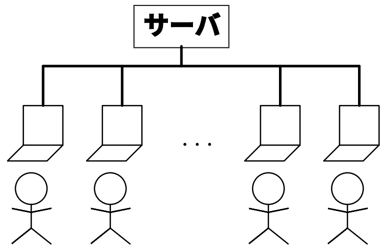

図 1　クライアント・サーバモデル

図 1 に示すクライアント・サーバモデルを考える．システムはサーバ 1 台とクライアント n 台を持つものとし，ジョブの集合を $U := \{1, \dots, m\}$, クライアントの集合を $V := \{1, \dots, n\}$ とする．利用者 $i \in V$ はクライアントマシン i を利用してジョブ $j = J(i) \in U$ を走らせているとする．ここで $J(i)$ は利用者 i のジョブの種類を表す．ジョブ j は対応する部分連鎖 j を持つとする．部分連鎖 j は，サーバと対応するノード j を交互に回る部分連鎖である．ジョブ j のサーバにおける平均サービス要求時間は $1/\mu_{\mathrm{s}j}$, クライアントにおける平均の思考時間は $1/\mu_{\mathrm{c}j}$ とする．

一つめのモデルとして，$n = m$ の場合を考える．このとき，利用者 i とそのジョブ $j = J(i)$ は同一視できる．このとき $\alpha_j := \mu_{\mathrm{c}j}/\mu_{\mathrm{s}j}$ とおくと，システムの定常分布は

$$P_{\mathrm{S}}(X) = \frac{|X|!}{G_{\mathrm{S}}} \prod_{j \in X} \alpha_j,$$
$$G_{\mathrm{S}} = \sum_{X \subset U} |X|! \prod_{j \in X} \alpha_j$$

と表される．

二つめのモデルは $n > m$ の場合であり，サーバに滞在するジョブ数ベクトルを $\boldsymbol{x} \in \mathbb{N}^m$ とすると，定常分布は

$$P_{\mathrm{M}}(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|!}{G_{\mathrm{M}}} \prod_{j=1}^{m} \frac{\alpha_j^{x_j}}{x_j!(n_j - x_j)!},$$
$$G_{\mathrm{M}} = \sum_{\boldsymbol{x} \in \Omega_{\mathrm{M}}} \|\boldsymbol{x}\|! \prod_{j=1}^{m} \frac{\alpha_j^{x_j}}{x_j!(n_j - x_j)!}$$

と表される．なお，$\|\boldsymbol{x}\| = x_1 + x_2 + \cdots + x_m$ と定義した．

このような形を積形式解と呼び，G_{S} や G_{M} を正規化定数と呼ぶ．システムの定常分布が積形式解で表される場合，その正規化定数の計算はシステムの性能評価において重要な部分を担う．一例として，正規化定数が計算で

きれば，ただちにシステムのスループットを計算することができる．スループットとは，単位時間当たりにシステムが処理を行ったトランザクションあるいはジョブの件数のことを言い，システム効率に関する性能評価指標として用いられている．

2.2 CFTP アルゴリズム

Coupling From The Past (CFTP) アルゴリズムは，1996 年に Propp と Wilson によって考案された．このアルゴリズムは確率的に停止し，そのときマルコフ連鎖の定常分布に厳密に従うサンプルを返す．

標準 CFTP アルゴリズムは停止したとき，マルコフ連鎖の定常分布に厳密に従って状態を出力する．しかしこの標準アルゴリズムでは状態空間中のすべての状態に対して操作を行わねばならず，現実的に実行が困難である．ただし，マルコフ連鎖 $\mathcal{M} : x \mapsto x'$ の更新関数表現 $x' = \phi(x, \Lambda)$ ($\Lambda \in [0, 1)$, 一様乱数) が以下に定義する単調性を持つとき，最大元および最小元に対してのみ操作を行えばよく，現実的に実行可能となる．

定義 1 (**単調な更新関数**). 状態空間 Ω 中に半順序関係 "$\succeq$" が存在し，「唯一の最大元 $x_{\max}$ および唯一の最小元 $x_{\min}$ が存在」し，かつ「任意の推移は半順序を保存する．すなわち，任意の $x \succeq y$ なる $x, y \in \Omega$ について，$\phi(x, \lambda) \succeq \phi(y, \lambda)$ が成り立つ」とき，ϕ は単調な更新関数であるという．

3 各連鎖の客数が 1 のクライアント・サーバモデル

3.1 正規化定数の計算法

正規化定数の計算は, 変数 x を持つ多項式 $(x+\alpha_1)(x+\alpha_2)\cdots(x+\alpha_m)$ を展開することで，客の人数 m に対し $\mathrm{O}(m^2)$ の時間複雑度で行うことができる．

3.2 定常分布からの完璧サンプリング法の構成

更新関数 ϕ_{S} によって表されるマルコフ連鎖 $\mathcal{M}_{\mathrm{S}}$ を，以下のように新たに構成する．

更新関数 $\phi_{\mathrm{S}} : 2^U \times U \times [0, 1) \to 2^U$ を

$$\phi_{\mathrm{S}}(X, j, \lambda) = \begin{cases} X \cup \{j\} & (\lambda < \theta_{jX}) \\ X \setminus \{j\} & (\text{otherwise}) \end{cases}$$

と定義する．なお，$\theta_{jX} := \frac{|X \cup \{j\}|\alpha_j}{|X \cup \{j\}|\alpha_j + 1}$ とする．これによりマルコフ連鎖 $\mathcal{M}_{\mathrm{S}} : X \mapsto X'$ を，それぞれ独立に一様ランダムに選んだ $j \in U$ および $\Lambda \in [0, 1)$ によって $X' = \phi_{\mathrm{S}}(X, j, \Lambda)$ と定義する．$\mathcal{M}_{\mathrm{S}}$ の定常分布はシステムの定常分布に等しい．

本研究では，ϕ_{S} が単調性を持つことを示した．

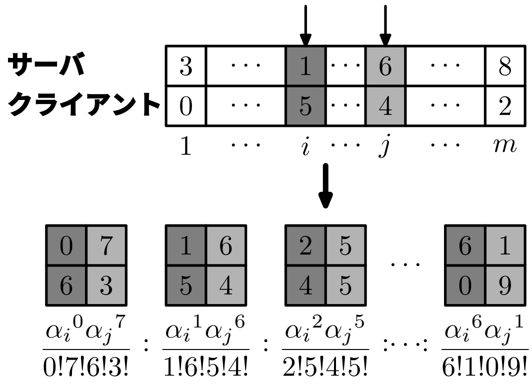

図 2 マルコフ連鎖 $\mathcal{M}_{\mathrm{M}}^h$

3.3 計算機実験

CFTP アルゴリズムの実行時間の指標となる coalescence time T_* について，その期待値を計算機実験によって計測した．$E[T_*]$ はデータの範囲内では m に対し指数関数的に増大することはなく，$\mathrm{O}(m^{1.5})$ 程度で押さえられると予想されるが，α_j の値にも依存し必ずしも両対数グラフ上できれいな直線には載らない．

4 各連鎖の客数が一般のクライアント・サーバモデル

正規化定数 G_{M} を求めるため，

$$G_{\mathrm{M}} = \sum_{h=0}^{n} h! G_{\mathrm{M}}^h,$$

$$G_{\mathrm{M}}^h = \sum_{\|\boldsymbol{x}\|=h} \prod_{j=1}^{m} \frac{\alpha_j^{x_j}}{x_j!(n_j - x_j)!}$$

と分解して，G_{M}^h $(h = 0, 1, \ldots, n)$ を求めることにする．そのため，各 h について，もとのネットワークの定常分布に対してサーバに滞在するジョブ数が h であるという条件付き分布を定常分布に持つマルコフ連鎖 $\mathcal{M}_{\mathrm{M}}^h$ を構成し，$\mathcal{M}_{\mathrm{M}}^h$ の定常分布からの近似サンプリングを行う．

4.1 マルコフ連鎖の構成

マルコフ連鎖 $\mathcal{M}_{\mathrm{M}}^h$ は，一様ランダムに選ばれたジョブ i, j についてのみ，サーバに滞在するジョブ数を変更するものである．その際，$x_i + x_j$ は一定に保たれる．概要を図 2 に示す．

4.2 Mixing time

以下の命題を示した．なお，mixing time $\tau(\varepsilon)$ は，マルコフ連鎖モンテカルロ法において，定常分布からの誤差が ε となるのにかかる推移回数を意味する．

命題 2. マルコフ連鎖 $\mathcal{M}_{\mathrm{M}}^h$ の mixing time $\tau_{\mathrm{M}}^h(\varepsilon)$ は，$\tau_{\mathrm{M}}^h(\varepsilon) \le \frac{m(m-1)}{2} \ln \frac{n}{\varepsilon}$ をみたす．