

一类 Fork-Join 排队系统的分析*

徐学雷 郑大钟

(清华大学自动化系·北京,100084)

摘要: 本文研究了一类具有有限排队空间且其到达率和服务率均依赖于状态的 Fork-Join 排队系统,给出了稳态概率和任务等待时间各阶矩的计算方法,并用仿真检验算法的正确性.

关键词: Fork-Join 排队系统; 建模与分析; 稳态概率; 等待时间

1 引言

考察图 1 所示的具有两个并联服务台的 Fork-Join 排队系统,每个作业到达 Fork 点立即分解成两个任务,分别送给服务台 1 和 2,使得每一服务台能且只能获得该作业的一个任务.各服务台按 FIFO 规则进行服务,且完成服务的任务总能输出,即各服务台后的存储容量足够大使得不会因不能输出而阻塞服务台;当某一作业的两个任务均完成服务时,该作业即在 Join 点重新组合起来,离开系统.称这种其到达和离开均具同步性的系统为 Fork-Join 系统,它在装配生产及并行计算中有广泛应用.

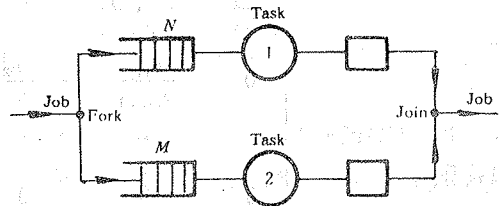


图 1 具有两个服务台的 Fork-Join 系统

对排队空间无限, Poisson 输入及指数分布服务时间的上述系统, [1]得到了稳态概率的母函数. [2]给出了任务等待时间联合分布的二维 Laplace 变换式. [3, 4]研究了一类其到达率和服务率均依赖于状态的单服务台系统.

本文将研究各排队空间有限且输入率和服务率均依赖于系统状态的上述 Fork-Join 排队系统.

2 稳态概率分布

设第一和第二队列的排队空间分别为 N 和 M , 若第一和第二队列中分别有 i 和 j 个任务(包括正在接受服务的任务), 则称系统处于 (i, j) 状态, 其状态空间为 $\{(i, j): 0 \leq i \leq N+1, 0 \leq j \leq M+1\}$. 若系统状态为 $(N+1, j)$ 或 $(i, M+1)$, 则不会有新的作业进入系统, 因此可设系统处于 (i, j) 状态时的输入率为 λ_{ij} , 第一及第二服务台的服务率分别为 μ_{1j} 和 μ_{2i} . 显然有 $\lambda_{i, M+1} = \lambda_{N+1, j} = 0, \mu_{0, j} = 0, \mu_{i, 0} = 0$.

令 t 时刻系统处于 (i, j) 状态的概率为 $P_{ij}(t)$, 由于系统具有不可约的有限 Markov 性, 因而极限 $\lim_{t \rightarrow \infty} P_{ij}(t) = P_{ij}$ 存在, 且 $\lim_{t \rightarrow \infty} \frac{dP_{ij}(t)}{dt} = 0$. 则可建立稳态概率分布 $\{P_{ij}\}$ 的平衡方程

* 国家自然科学基金和863自动化领域 CIMS 主题资助项目.

本文于1992年4月14日收到. 1993年12月7日收到修改稿.

$$a_{ij}P_{ij} = \mu_{i+1,j}P_{i+1,j} + \gamma_{i,j+1}P_{i,j+1} + \lambda_{i-1,j-1}P_{i-1,j-1}, \quad 0 \leq i \leq N, \quad 0 \leq j \leq M, \quad (1)$$

$$a_{N+1,j}P_{N+1,j} = \gamma_{N+1,j+1}P_{N+1,j+1} + \lambda_{N,j-1}P_{N,j-1}, \quad 0 \leq j \leq M, \quad (2)$$

$$a_{i,M+1}P_{i,M+1} = \mu_{i+1,M+1}P_{i+1,M+1} + \lambda_{i-1,M}P_{i-1,M}, \quad 0 \leq i \leq N, \quad (3)$$

$$a_{N+1,M+1}P_{N+1,M+1} = \lambda_{N,M}P_{N,M}. \quad (4)$$

其中

$$a_{ij} = \mu_{ij} + \gamma_{ij} + \lambda_{ij}.$$

由于方程(4)可由方程组(1)~(3)推出,因此在求解 $\{P_{ij}\}$ 的过程中将(4)去掉,并将方程(1)~(3)与概率完全性方程

$$\sum_{j=0}^{M+1} \sum_{i=0}^{N+1} P_{ij} = 1 \quad (5)$$

联立即可解得 $\{P_{ij}\}$.下面将用分块矩阵方法求解 $\{P_{ij}\}$,它将大大地减少计算量.令 $\mathcal{D}_j = \{P_{0j}, \dots, P_{N+1,j}\}^T, j=0, 1, \dots, M+1$,其中 A^T 表示矩阵 A 的转置;再令

$$\mathcal{A}_j = \begin{bmatrix} a_{0,j} & -\mu_{1,j} & 0 & \dots & 0 & 0 \\ \gamma_{0,j+1} & \gamma_{0,j+1} & & & & \\ 0 & a_{1,j} & -\mu_{2,j} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & & a_{N,j} & -\mu_{N+1,j} \\ 0 & 0 & 0 & \dots & \gamma_{N,j+1} & \gamma_{N,j+1} \\ 0 & 0 & 0 & \dots & 0 & a_{N+1,j} \\ & & & & & \gamma_{N+1,j+1} \end{bmatrix}_{(N+2) \times (N+2)}$$

$$\mathcal{B}_j = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ -\lambda_{0,j-1} & 0 & \dots & 0 & 0 \\ \gamma_{1,j+1} & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -\lambda_{N,j-1} & 0 \\ & & & \gamma_{N+1,j+1} & \end{bmatrix}_{(N+2) \times (N+2)}$$

则方程(1),(2)可表为如下分块矩阵形式

$$\begin{bmatrix} -I \\ \mathcal{A}_1 & -I & & 0 \\ \mathcal{B}_2 & \mathcal{A}_2 & -I & \\ & \ddots & \ddots & \ddots \\ 0 & \mathcal{B}_N & \mathcal{A}_N & -I \end{bmatrix} \begin{bmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \mathcal{D}_3 \\ \vdots \\ \mathcal{D}_{M+1} \end{bmatrix} = \begin{bmatrix} -\mathcal{A}_0 \mathcal{D}_1 \\ -\mathcal{B}_1 \mathcal{D}_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (6)$$

其中 I 为 $(N+2) \times (N+2)$ 阶单位矩阵.由(6)知:其左端的系数矩阵是可逆的,利用分块矩阵求逆可得:

$$\mathcal{D}_j = D_j \mathcal{D}_0, \quad j = 1, 2, \dots, M+1. \quad (7)$$

其中 $D_j = Q_{1,j-1} \mathcal{A}_0 + Q_{2,j-1} \mathcal{B}_1, Q_{ij} = \mathcal{A}_j Q_{i,j-1} + \mathcal{B}_j Q_{i,j-2}, Q_{i,i-2} = 0, Q_{i,i-1} = I, Q_{ii} = \mathcal{A}_i, i = 1, 2; j = 1, 2, \dots, M+1$.

将(7)代入(3)和(5)式,并表成矩阵形式得

$$\begin{cases} W\mathcal{D}_0 = 0, \\ [1, 1, \dots, 1](I + \sum_{j=1}^{M+1} D_j)\mathcal{D}_0 = 1. \end{cases} \quad (8)$$

其中 W 为 $(N+1) \times (N+2)$ 阶矩阵, 其 (i, j) 位置上元素为

$$w_{ij} = a_{i, M+1}d_{ij}^{M+1} - \mu_{i+1, M+1}d_{i+1, j}^{M+1} - \lambda_{i-1, N}d_{i-1, j}^M, \quad i = 0, 1, \dots, N, \quad j = 0, 1, \dots, N+1.$$

而 d_{ij}^k 为 D_k 的第 (i, j) 位置上的元素.

于是由 (8) 解出 \mathcal{D}_0 , 并代入 (7) 即得 $\mathcal{D}_j (j=1, \dots, M+1)$, 至此即获得稳态概率分布 $\{P_{ij}\}$.

在获得 $\{P_{ij}\}$ 的基础上, 若设第 i 队列的稳态队长为 $N_i (i=1, 2)$, 则 $P(N_1=i) = \sum_{j=0}^{M+1} P_{ij}$, $P(N_2=i) = \sum_{j=0}^{M+1} P_{ij}$, 所以, 平均稳态队长为: $E(N_1) = \sum_{i=0}^{N+1} \sum_{j=0}^{M+1} iP_{ij}$, $E(N_2) = \sum_{i=0}^{N+1} \sum_{j=0}^{M+1} jP_{ij}$.

3 等待时间的矩

任务等待时间即是从任务到达时起到它开始接受服务时这段时间, 本节讨论各队列中任务等待时间的各阶原点矩.

设从系统处于 (i, j) 状态开始, 第一(或二)服务台连续完成 k 次服务所需时间为 s_k^i (或 t_k^j), 其分布密度函数为 $h_k^i(t)$ (或 $g_k^j(t)$), 由任务服务时间的无记忆性, 且其剩余服务时间服从 $s_{k+1, j+1}^i$ 或 $s_{k, j-1}^i$ 的分布, 则可得 $h_k^i(t)$ 满足递推方程

$$\begin{cases} h_k^i(t) = \mu_{ij}e^{-a_{ij}t} + \lambda_{ij}e^{-a_{ij}t} * h_{k+1, j+1}^i(t) + \gamma_{ij}e^{-a_{ij}t} * h_{k, j-1}^i(t), \\ h_k^j(t) = \mu_{ij}e^{-a_{ij}t} * h_{k-1, j}^j(t) + \lambda_{ij}e^{-a_{ij}t} * h_{k+1, j+1}^j(t) + \gamma_{ij}e^{-a_{ij}t} * h_{k, j-1}^j(t), \end{cases} \quad (9)$$

$i = 1, 2, \dots, N+1; \quad j = 0, 1, \dots, M+1; \quad k = 2, \dots, i.$

其中 $*$ 表示卷积, 且当 $i > N+1, j < 0$ 或 $j > M+1$ 时有 $h_k^i(t) = 0$.

取 $H_i^q = \int_0^\infty t^q h_k^i(t) dt$ 为 s_k^i 的 q 阶原点矩. 令 $\mathcal{H}_i^q = (H_{i0}^q, H_{i1}^q, \dots, H_{i, M+1}^q)^T$, 以及

$$R_i = \begin{bmatrix} a_{i0} & & & & & \\ -\gamma_{i1} & a_{i1} & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ 0 & & & & -\gamma_{i, M+1} & a_{i, M+1} \end{bmatrix}, \quad M_i = \begin{bmatrix} 0 & \lambda_{i0} & & & & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \lambda_{iM} & \\ & & & & & 0 \end{bmatrix},$$

$$A_i = \begin{bmatrix} \mu_{i0} & & & & & \\ & \mu_{i1} & & & & \\ & & \ddots & & & \\ 0 & & & \ddots & & \\ & & & & & \mu_{i, M+1} \end{bmatrix},$$

则将上述递推方程组 (9) 两端同时作 Laplace 变换, 并在原点处求 q 阶导数, 经整理可得如下矩阵形式的递推方程组:

$$\begin{aligned} R_i \mathcal{H}_i^q &= M_i \mathcal{H}_{i+1}^q + q \mathcal{H}_i^{q-1}, \\ R_i \mathcal{H}_i^q &= M_i \mathcal{H}_{i+1}^q + A_i \mathcal{H}_{i-1}^{q-1} + q \mathcal{H}_i^{q-1}, \quad i = 1, 2, \dots, N+1, \quad k = 2, \dots, i. \end{aligned}$$

当 $i > N+1$ 时, $\mathcal{H}_i^q = 0$.

若系统处于 (i, j) 状态时, 有一作业进入, 则该作业的任务只有等到第一(二)队列中完成 $i(j)$ 次服务后, 才能开始接受服务. 设作业看到系统处于 (i, j) 状态时进入的概率为 Π_{ij} , 则

$$\Pi_{ij} = \lambda_{ij} P_{ij} / \left(\sum_{i=0}^N \sum_{j=0}^M \lambda_{ij} P_{ij} \right), \quad i = 0, 1, \dots, N; j = 0, 1, \dots, M.$$

因此, 队列 1 中任务等待时间 W_1 的 q 阶原点矩为

$$\bar{W}_1^q = \sum_{i=0}^N \sum_{j=0}^M \Pi_{ij} H_{ij}^q, \quad q = 1, 2, \dots.$$

同理, 若取 $G_j^q = \int_0^\infty t^q g_j^*(t) dt$ 为 t_j 的 q 阶原点矩, $\mathcal{G}_j^q = (G_{0j}^q, G_{1j}^q, \dots, G_{N+1,j}^q)^T$, 以及

$$R_j = \begin{bmatrix} a_{0j} & & & & \\ -\mu_{1j} & a_{1j} & & & 0 \\ & \ddots & \ddots & & \\ 0 & & & -\mu_{N+1,j} & a_{N+1,j} \end{bmatrix}, \quad M_j = \begin{bmatrix} 0 & \lambda_{0j} & & & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & \lambda_{Nj} & \\ & & & & 0 \end{bmatrix},$$

$$A_j = \begin{bmatrix} \gamma_{0j} & & & & 0 \\ & \gamma_{1j} & & & \\ 0 & & \ddots & & \\ & & & \gamma_{N+1,j} & \end{bmatrix}.$$

则 \mathcal{G}_j^q 满足如下递推方程

$$R_j \mathcal{G}_j^q = M_j \mathcal{G}_{j+1}^q + q \mathcal{G}_j^{q-1},$$

$$R_j \mathcal{G}_j^q = M_j \mathcal{G}_{j+1}^q + A_j \mathcal{G}_{j-1}^q + q \mathcal{G}_j^{q-1}, \quad j = 1, 2, \dots, M+1, \quad k = 2, \dots, j.$$

队列 2 中任务等待时间 W_2 的 q 阶原点矩为

$$\bar{W}_2^q = \sum_{i=0}^N \sum_{j=0}^M \Pi_{ij} H_{ij}^q, \quad q = 1, 2, \dots.$$

4 算例及仿真结果

当 $N=M=5$ 时, 用本文的算法对 $(\lambda_{ij}, \mu_{ij}, \gamma_{ij}) = (\lambda, 1, 0.5)$ 系统作了数值计算, 并用 C 语言进行了仿真, 其结果如表 1 和表 2 所示.

表 1 理论计算值

λ	0.2	0.4	0.6	0.8	1.0
\bar{W}_1	0.247094	0.552240	0.765554	0.866560	0.911023
\bar{W}_2	1.283689	3.728930	6.018684	7.393401	8.152955

表 2 仿真结果(取 10000 个时间单位)

λ	0.2	0.4	0.6	0.8	1.0
\bar{W}_1	0.252991	0.589142	0.780981	0.912897	0.985576
\bar{W}_2	1.309416	4.031234	6.166146	7.499261	8.213132

从上面两个表可以看出: 本文的算法与仿真结果的误差小于 10%, 因此本文的算法是有效的.

5 结束语

本文研究了一类状态依赖的 Fork-Join 排队系统,给出了稳态概率及等待时间各阶矩的算法;而平均作业响应时间需进一步研究.

参 考 文 献

- [1] Flatto, L. and Hahn, S. Two Parallel Queues Created by Arrivals with Two Demands(1). *SIAM J. Appl. Math.*, 1984, 44(5):1041—1053
- [2] Zhang, Z. S. Analytical Results for Waiting Time and System Size Distributions in Two Parallel Queuing Systems. *SIAM J. Appl. Math.*, 1990, 50(4):1176—1193
- [3] Van Dooren, E. The Transient State Probabilities for A Queuing Model Where Potential Customers Are Discouraged by Queue Length. *J. Appl. Prob.*, 1981, 18(2):499—509
- [4] Conolly, B. and Chan, J. Generalized Birth and Death Queuing Processes, Recent Results. *Adv. Appl. Prob.*, 1977, 9(1):125—140

Analysis of a Class of Fork-Join Queuing System

XU Xuelei and ZHENG Dazhong

(Department of Automation, Tsinghua University • Beijing, 100084, PRC)

Abstract: A class of fork-join queuing systems with finite queuing spaces and state-dependent arrivals and service rates is considered. A recursive set of equations for the steady-state probabilities is obtained. The moments of the task waiting time in each queue are derived. The results are checked by simulation.

Key words: fork-join queuing system; modelling and analysis; steady-state probability; waiting time

本文作者简介

徐学雷 1963年生. 1983年毕业于四川大学, 1986年在西北工业大学获得硕士学位, 并留校任教, 1988年被评为讲师. 现在清华大学自动化系攻读博士学位. 主要从事离散事件动态系统的研究工作.

郑大钟 1935年生. 1959年毕业于清华大学自动控制系. 现为清华大学自动化系教授. 主要研究领域有线性系统理论, 最优控制, 大系统分散控制, 鲁棒控制, 离散事件动态系统等.