

具有中途准入机制和多重休假的离散时间 $GI/Geom^{(a,b)}/1/N$ 早到排队系统*

余 妙

(四川师范大学数学与软件科学学院, 成都 610068)

(四川理工学院理学院, 自贡 643000)

唐应辉†

(四川师范大学数学与软件科学学院, 成都 610068)

(E-mail: tangyh@uestc.edu.cn)

付永红

(清华大学计算机科学与技术系, 北京 100084)

摘 要 运用补充变量方法和嵌入 Markov 链方法讨论了一个具有批量服务中途准入机制的离散时间多重休假排队系统. 利用一种有效的数值迭代算法获得了系统中三种时刻的队长分布. 进一步, 使用不同时刻的队长分布, 通过数值化方法研究了系统参数对阻塞概率, 批量服务中途准入概率, 顾客在缓冲空间中平均等待时间等几类重要性能指标的影响.

关键词 离散时间排队; 批量服务中途准入机制; 补充变量方法; 嵌入 Markov 链

MR(2000) 主题分类 60K25; 90B22

中图分类 O226

1 引言

批量服务排队模型是经典的逐个服务排队模型在顾客服务环节的一种推广和拓展, 它已被广泛应用于大规模生产制造系统的数学建模和运作管理中^[1]. 受服务员服务能力所限, 批量服务排队系统中通常都设定了一个服务批量的上限值 b , 当系统中顾客个

本文 2010 年 8 月 28 日收到. 2011 年 5 月 4 日收到修改稿.

* 国家自然科学基金 (70871084), 教育部高校博士点基金项目 (200806360001) 以及四川省教育厅自然科学基金重点 (10ZA136) 资助项目.

† 通信作者.

数超过 b 时, 超出部分的顾客将在队列中等候下一次的服务. Chaudhry 和 Templeton^[2], Vijaya Laxmi 和 Gupta^[3] 先后利用补充变量和嵌入 Markov 链方法研究了用连续时间刻画的 Poisson 输入 $M/G^b/1$ 和一般更新输入 $GI/M^b/1/N$ 排队系统. 作为上述工作的进一步发展 Sivasamy^[4] 和 Samanta^[5] 等人从节约系统运行成本的角度出发又提出了一种广义的 (a, b) 批量服务策略, 即当系统中至少有 a 个顾客存在时服务才会进行且每次服务批量的上限值为 b . 值得注意的是, (a, b) 批量服务策略中的阈值 a 与 N 策略中的阈值 N 有本质上的区别. 在 N 策略排队中服务一旦启动服务员将在系统空竭后才进入休眠或休假状态, 而这里的阈值 a 表示只要在完成一次服务后服务员发现系统中顾客个数小于 a 则立即进入休眠或休假状态, 因此有时这个系统会是一个非空竭服务系统. 在上述 (a, b) 批量服务策略中, 服务启动后即使当前服务批量未达到上限值 b , 其后到达的顾客也不能加入其中接受服务. 前不久, Gupta 和 Sikdar^[6] 将这种批量服务策略与服务员单重休假机制相结合, 详细研究了缓冲空间有限的 $M/G/1$ 排队系统. 事实上, 这样的 (a, b) 批量服务策略在实际应用中不尽如人意之处, 我们注意到当正在服务的顾客批量大于 a 而小于 b 时服务员并未尽其所能在工作, 也就是说系统此时有闲置富余的服务能力存在, 而在生产实践中系统管理者, 生产经营商总是希望“物尽其能”在最大程度上给自己带来丰厚高额的收益. 因此当他们发现服务员未满负荷工作时通常会允许后到的顾客中途加入正在进行的批量服务中. 例如, 私人承包的长途客运线路, 考虑到燃油、公路过路费等固有成本, 司机总是要等到乘客人数达到一定量后才能发班, 但可能此时车上乘客并未达到客车准载人数的上限, 所以中途若有乘客加入时, 司机会让他上车以求在营运过程中实现收益的最大化. 针对 (a, b) 批量服务策略的不足, 最近 Arumuganathan 和 Rathinasamy^[7], Vijaya Laxmi, Goswami 和 Yesuf^[8] 等人研究了具有批量服务中途准入机制的连续时间 $M/G/1$ 和 $GI/M/1/N$ 排队系统, 并且获得了系统的诸多重要排队性能指标. 然而到目前为止关于具有批量服务中途准入机制的离散时间排队系统的研究并不多见, 仅有 Goswami, Mohanty 和 Samanta^[9] 等人利用 Kolmogorov 方程和迭代技术研究过一个简单的离散时间马氏排队系统. 本文的目的就是要在前人的工作基础上将批量服务中途准入策略的研究推广到非马氏的离散时间休假排队系统中, 从而为这一策略更好的在数字通信和计算机网络性能评估中应用提供一些必要的理论依据.

2 模型描述与符号说明

本文研究具有中途准入机制和多重休假的离散时间 $GI/Geom^{(a,b)}/1/N$ 早到达排队系统, 系统的具体描述如下:

(1) 按照 EAS (Early arrival system) 入口协议规定^[10,11], 潜在顾客的到达发生于时隙首端 (t, t^+) 处, 服务的结束只能发生在时隙的末端 (t^-, t) 处 (见图 1);

(2) 相邻两顾客的到达间隔时间服从取值为正整数的一般离散概率分布, 记为 τ , 其概率分布律为 $\text{Prob}\{\tau = k\} = g_k, k \geq 1$, 且 $E[\tau] = 1/\lambda$, 概率母函数为 $G(z) = \sum_{k=0}^{\infty} g_k z^k$ ($g_0 = 0, |z| \leq 1$);

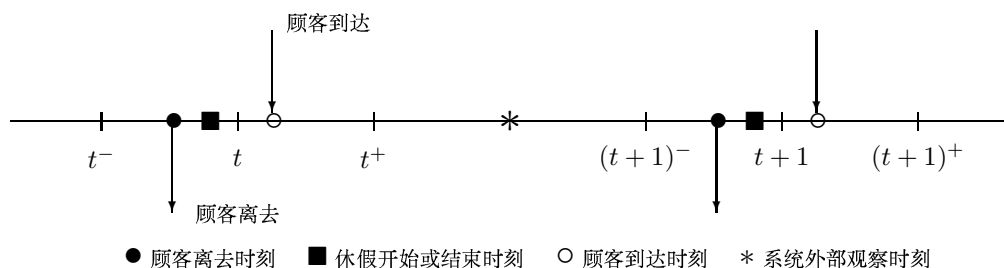


图 1 EAS 系统的几个特殊时刻

(3) 服务员按照中途准入的 (a, b) 批量服务策略向顾客提供服务, 每批顾客的服务时间 S_l ($l = 1, 2, \dots$) 服从参数为 μ 的几何分布, 即 $\text{Prob}\{S_l = n\} = \mu\bar{\mu}^{n-1}$, $0 < \mu < 1$, $\bar{\mu} = 1 - \mu$, $n \geq 1$. 这里, 中途准入的 (a, b) 批量服务策略具体定义如下: 当一次批量服务结束后, 如果缓冲队列中的顾客数小于 a , 则服务员立即进入随机长度的休假期 V , 直至某次休假结束发现缓冲队列中等待的顾客人数大于等于 a 后再重新进入服务状态; 若服务员当前正在服务的顾客数量大于等于 a 而小于 b , 则此时新到达的顾客不在缓冲空间中等待, 直接加入当前的批量服务中接受服务; 若当前正在接受服务的顾客数量已达批量上限值 b , 且在系统缓冲空间中等待的顾客数量小于 N , 则新到顾客进入缓冲空间等候服务; 否则当缓冲空间无富余容量时新到顾客受阻消失. 休假时间 V 服从参数为 θ 的几何分布, 即 $\text{Prob}\{V = n\} = \theta\bar{\theta}^{n-1}$, $0 < \theta < 1$, $\bar{\theta} = 1 - \theta$, $n \geq 1$;

(4) 排队系统容量有限, 包括正在接受服务的顾客在内系统的最大容量为 $b + N$;

(5) 顾客达到过程, 服务过程及其服务员休假过程相互独立.

系统在时刻 t 的状态可以由如下四个随机变量加以描述:

$N_q(t)$: t 时刻系统缓冲空间中顾客人数;

$N_s(t)$: t 时刻正在接受服务的顾客人数;

$U(t)$: t 时刻顾客的剩余到达间隔时间;

$$Y(t) = \begin{cases} 0, & t \text{ 时刻服务员处于休假状态;} \\ 1, & t \text{ 时刻服务员繁忙且此时正在服务的顾客人数未达到批量上限 } b; \\ 2, & t \text{ 时刻服务员繁忙且此时正在服务的顾客人数已达批量上限 } b. \end{cases}$$

由 [12] 易知 $\{N_s(t), N_q(t), U(t), Y(t)\}$ 构成了一个高维的离散时间 Markov 过程, 系统在时刻 t 的全部状态概率可定义如下:

$$P_{i,0}(t, u) = \text{Prob}\{N_q(t) = i, U(t) = u, Y(t) = 0\}, \quad i = 0, \dots, N, \quad u \geq 0;$$

$$P_{i,1}(t, u) = \text{Prob}\{N_s(t) = i, U(t) = u, Y(t) = 1\}, \quad i = a, \dots, b-1, \quad u \geq 0;$$

$$P_{i,2}(t, u) = \text{Prob}\{N_q(t) = i, U(t) = u, Y(t) = 2\}, \quad i = 0, \dots, N, \quad u \geq 0.$$

注 1 当 $Y(t) = 0$ 即服务员处于休假状态时 $N_s(t) \equiv 0$; 当 $Y(t) = 1$ 即服务员处于繁忙状态且当前正在接受服务的顾客人数未达到批量上限值 b 时 $N_q \equiv 0$; 当 $Y(t) = 2$ 即服务员处于繁忙状态且当前正在接受服务的顾客人数已达到批量上限值 b 时 $N_s \equiv b$. 为

使状态概率的表达形式更加简洁,同时保证系统状态个数不发生变化,在上述三个状态概率的定义中我们分别省略了 $N_s(t) \equiv 0$, $N_q \equiv 0$ 和 $N_s \equiv b$ 这三个分量.

进一步,在全文的分析中我们还会用到如下一些符号:

$$\begin{aligned} P_{i,0}(u) &= \lim_{t \rightarrow \infty} P_{i,0}(t, u), \quad i = 0, \dots, N, \quad u \geq 0; \\ P_{i,1}(u) &= \lim_{t \rightarrow \infty} P_{i,1}(t, u), \quad i = a, \dots, b-1, \quad u \geq 0; \\ P_{i,2}(u) &= \lim_{t \rightarrow \infty} P_{i,2}(t, u), \quad i = 0, \dots, N, \quad u \geq 0; \\ P_{i,0} &= \lim_{t \rightarrow \infty} \text{Prob}\{N_q(t) = i, Y(t) = 0\}, \quad i = 0, \dots, N; \\ P_{i,1} &= \lim_{t \rightarrow \infty} \text{Prob}\{N_s(t) = i, Y(t) = 1\}, \quad i = a, \dots, b-1; \\ P_{i,2} &= \lim_{t \rightarrow \infty} \text{Prob}\{N_q(t) = i, Y(t) = 2\}, \quad i = 0, \dots, N; \\ P_{i,0}^*(z) &= \sum_{u=0}^{\infty} P_{i,0}(u) z^u, \quad |z| \leq 1, \quad i = 0, \dots, N; \\ P_{i,1}^*(z) &= \sum_{u=0}^{\infty} P_{i,1}(u) z^u, \quad |z| \leq 1, \quad i = a, \dots, b-1; \\ P_{i,2}^*(z) &= \sum_{u=0}^{\infty} P_{i,2}(u) z^u, \quad |z| \leq 1, \quad i = 0, \dots, N. \end{aligned}$$

3 三种时刻系统稳态队长分布的计算方法

本节将讨论稳态情形下任意时刻,顾客到达前夕和系统外部观察时刻的三种队长分布.首先使用离散补充变量方法和 Bayes 公式获得刻画任意时刻与顾客到达前夕队长分布的一组迭代关系,其次结合 EAS 系统状态的演化进程给出任意时刻与系统外部观察时刻队长分布的迭代关系,最后利用嵌入 Markov 链方法所得到的转移概率矩阵求解顾客到达前夕的稳态队长分布,并通过上述所得到的一系列迭代公式获得系统任意时刻和外部观察时刻的队长分布.

3.1 系统任意时刻与顾客到达前夕队长分布的迭代关系

考虑系统在时刻 t 和 $t+1$ 处的状态转移情况,经过简单的概率分析容易建立起各状态概率之间的一组差分方程

$$P_{0,0}(t+1, u-1) = P_{0,0}(t, u) + \mu \sum_{i=a}^{b-1} P_{i,1}(t, u) + g_u \mu \sum_{i=a}^{b-1} P_{i,1}(t, 0) + \mu P_{0,2}(t, u), \quad (1)$$

$$\begin{aligned} P_{i,0}(t+1, u-1) &= P_{i,0}(t, u) + g_u P_{i-1,0}(t, 0) \\ &\quad + \mu P_{i,2}(t, u) + g_u \mu P_{i-1,2}(t, 0), \quad i = 1, \dots, a-1, \end{aligned} \quad (2)$$

$$P_{i,0}(t+1, u-1) = \bar{\theta} P_{i,0}(t, u) + g_u \bar{\theta} P_{i-1,0}(t, 0), \quad i = a, \dots, N-1, \quad (3)$$

$$P_{N,0}(t+1, u-1) = \bar{\theta} P_{N,0}(t, u) + g_u \bar{\theta} P_{N-1,0}(t, 0) + g_u \bar{\theta} P_{N,0}(t, 0), \quad (4)$$

$$\begin{aligned} P_{a,1}(t+1, u-1) &= \bar{\mu} P_{a,1}(t, u) + \mu P_{0,2}(t, u) + \theta P_{a,0}(t, u) \\ &\quad + g_u \theta P_{a-1,0}(t, 0) + g_u \mu P_{a-1,2}(t, 0), \end{aligned} \quad (5)$$

$$P_{i,1}(t+1, u-1) = \bar{\mu} P_{i,1}(t, u) + \mu P_{i,2}(t, u) + \theta P_{i,0}(t, u) + g_u \theta P_{i-1,0}(t, 0)$$

$$+ g_u \bar{\mu} P_{i-1,1}(t, 0) + g_u \mu P_{i-1,2}(t, 0), \quad i = a + 1, \dots, b - 1, \quad (6)$$

$$P_{0,2}(t + 1, u - 1) = \bar{\mu} P_{0,2}(t, u) + g_u \bar{\mu} P_{b-1,1}(t, 0) + \theta P_{b,0}(t, u) + g_u \theta P_{b-1,0}(t, 0) + \mu P_{b,2}(t, u) + g_u \mu P_{b-1,2}(t, 0), \quad (7)$$

$$P_{i,2}(t + 1, u - 1) = \bar{\mu} P_{i,2}(t, u) + \theta P_{i+b,0}(t, u) + g_u \theta P_{i+b-1,0}(t, 0) + \mu P_{i+b,2}(t, u) + g_u \mu P_{i+b-1,2}(t, 0) + g_u \bar{\mu} P_{i-1,2}(t, 0), \quad i = 1, \dots, N - b - 1, \quad (8)$$

$$P_{N-b,2}(t + 1, u - 1) = \bar{\mu} P_{N-b,2}(t, u) + \theta P_{N,0}(t, u) + g_u \theta P_{N-1,0}(t, 0) + g_u \theta P_{N,0}(t, 0) + \mu P_{N,2}(t, u) + g_u \mu P_{N-1,2}(t, 0) + g_u \mu P_{N,2}(t, 0) + g_u \bar{\mu} P_{N-b-1,2}(t, 0), \quad (9)$$

$$P_{i,2}(t + 1, u - 1) = \bar{\mu} P_{i,2}(t, u) + g_u \bar{\mu} P_{i-1,2}(t, 0), \quad i = N - b + 1, \dots, N - 1, \quad (10)$$

$$P_{N,2}(t + 1, u - 1) = \bar{\mu} P_{N,2}(t, u) + g_u \bar{\mu} P_{N-1,2}(t, 0) + g_u \bar{\mu} P_{N,2}(t, 0). \quad (11)$$

在 (1)–(11) 式中令 $t \rightarrow \infty$, 上述方程可简化为

$$P_{0,0}(u - 1) = P_{0,0}(u) + \mu \sum_{i=a}^{b-1} P_{i,1}(u) + g_u \mu \sum_{i=a}^{b-1} P_{i,1}(0) + \mu P_{0,2}(u), \quad (12)$$

$$P_{i,0}(u - 1) = P_{i,0}(u) + g_u P_{i-1,0}(0) + \mu P_{i,2}(u) + g_u \mu P_{i-1,2}(0), \quad i = 1, \dots, a - 1, \quad (13)$$

$$P_{i,0}(u - 1) = \bar{\theta} P_{i,0}(u) + g_u \bar{\theta} P_{i-1,0}(0), \quad i = a, \dots, N - 1, \quad (14)$$

$$P_{N,0}(u - 1) = \bar{\theta} P_{N,0}(u) + g_u \bar{\theta} P_{N-1,0}(0) + g_u \bar{\theta} P_{N,0}(0), \quad (15)$$

$$P_{a,1}(u - 1) = \bar{\mu} P_{a,1}(u) + \mu P_{a,2}(u) + \theta P_{a,0}(u) + g_u \theta P_{a-1,0}(0) + g_u \mu P_{a-1,2}(0), \quad (16)$$

$$P_{i,1}(u - 1) = \bar{\mu} P_{i,1}(u) + \mu P_{i,2}(u) + \theta P_{i,0}(u) + g_u \theta P_{i-1,0}(0) + g_u \bar{\mu} P_{i-1,1}(0) + g_u \mu P_{i-1,2}(0), \quad i = a + 1, \dots, b - 1, \quad (17)$$

$$P_{0,2}(u - 1) = \bar{\mu} P_{0,2}(u) + g_u \bar{\mu} P_{b-1,1}(0) + \theta P_{b,0}(u) + g_u \theta P_{b-1,0}(0) + \mu P_{b,2}(u) + g_u \mu P_{b-1,2}(0), \quad (18)$$

$$P_{i,2}(u - 1) = \bar{\mu} P_{i,2}(u) + \theta P_{i+b,0}(u) + g_u \theta P_{i+b-1,0}(0) + \mu P_{i+b,2}(u) + g_u \mu P_{i+b-1,2}(0) + g_u \bar{\mu} P_{i-1,2}(0), \quad i = 1, \dots, N - b - 1, \quad (19)$$

$$P_{N-b,2}(u - 1) = \bar{\mu} P_{N-b,2}(u) + \theta P_{N,0}(u) + g_u \theta P_{N,0}(0) + g_u \theta P_{N-1,0}(0) + \mu P_{N,2}(u) + g_u \mu P_{N-1,2}(0) + g_u \mu P_{N,2}(0) + g_u \bar{\mu} P_{N-b-1,2}(0), \quad (20)$$

$$P_{i,2}(u - 1) = \bar{\mu} P_{i,2}(u) + g_u \bar{\mu} P_{i-1,2}(0), \quad i = N - b + 1, \dots, N - 1, \quad (21)$$

$$P_{N,2}(u - 1) = \bar{\mu} P_{N,2}(u) + g_u \bar{\mu} P_{N-1,2}(0) + g_u \bar{\mu} P_{N,2}(0). \quad (22)$$

对 (12)–(22) 式左右两端同时施行 Z 变换后有:

$$(z - 1) P_{0,0}^*(z) = \mu \sum_{i=a}^{b-1} P_{i,1}^*(z) + G(z) \mu \sum_{i=a}^{b-1} P_{i,1}(0) + \mu P_{0,2}^*(z)$$

$$-P_{0,0}(0) - \mu \sum_{i=a}^{b-1} P_{i,1}(0) - \mu P_{0,2}(0), \quad (23)$$

$$(z-1)P_{i,0}^*(z) = G(z)P_{i-1,0}(0) + \mu P_{i,2}^*(z) + G(z)\mu P_{i-1,2}(0) - P_{i,0}(0) - \mu P_{i,2}(0), \quad i=1, \dots, a-1, \quad (24)$$

$$(z-\bar{\theta})P_{i,0}^*(z) = G(z)\bar{\theta}P_{i-1,0}(0) - \bar{\theta}P_{i,0}(0), \quad i=a, \dots, N-1, \quad (25)$$

$$(z-\bar{\theta})P_{N,0}^*(z) = G(z)\bar{\theta}P_{N-1,0}(0) + G(z)\bar{\theta}P_{N,0}(0) - \bar{\theta}P_{N,0}(0), \quad (26)$$

$$(z-\bar{\mu})P_{a,1}^*(z) = \mu P_{a,2}^*(z) + \theta P_{a,0}^*(z) + G(z)\theta P_{a-1,0}(0) + G(z)\mu P_{a-1,2}(0) - \bar{\mu}P_{a,1}(0) - \mu P_{a,2}(0) - \theta P_{a,0}(0), \quad (27)$$

$$(z-\bar{\mu})P_{i,1}^*(z) = \mu P_{i,2}^*(z) + \theta P_{i,0}^*(z) + G(z)\theta P_{i-1,0}(0) + G(z)\bar{\mu}P_{i-1,1}(0) + G(z)\mu P_{i-1,2}(0) - \bar{\mu}P_{i,1}(0) - \mu P_{i,2}(0) - \theta P_{i,0}(0), \quad i=a+1, \dots, b-1, \quad (28)$$

$$(z-\bar{\mu})P_{0,2}^*(z) = \theta P_{b,0}^*(z) + \mu P_{b,2}^*(z) + G(z)\bar{\mu}P_{b-1,1}(0) + G(z)\theta P_{b-1,0}(0) + G(z)\mu P_{b-1,2}(0) - \bar{\mu}P_{0,2}(0) - \theta P_{b,0}(0) - \mu P_{b,2}(0), \quad (29)$$

$$(z-\bar{\mu})P_{i,2}^*(z) = \theta P_{i+b,0}^*(z) + \mu P_{i+b,2}^*(z) + G(z)\theta P_{i+b-1,0}(0) + G(z)\mu P_{i+b-1,2}(0) + G(z)\bar{\mu}P_{i-1,2}(0) - \bar{\mu}P_{i,2}(0) - \theta P_{i+b,0}(0) - \mu P_{i+b,2}(0), \quad i=1, \dots, N-b-1, \quad (30)$$

$$(z-\bar{\mu})P_{N-b,2}^*(z) = \theta P_{N,0}^*(z) + \mu P_{N,2}^*(z) + G(z)\theta P_{N-1,0}(0) + G(z)\theta P_{N,0}(0) + G(z)\mu P_{N-1,2}(0) + G(z)\mu P_{N,2}(0) + G(z)\bar{\mu}P_{N-b-1,2}(0) - \bar{\mu}P_{N-b,2}(0) - \theta P_{N,0}(0) - \mu P_{N,2}(0), \quad (31)$$

$$(z-\bar{\mu})P_{i,2}^*(z) = G(z)\bar{\mu}P_{i-1,2}(0) - \bar{\mu}P_{i,2}(0), \quad i=N-b+1, \dots, N-1, \quad (32)$$

$$(z-\bar{\mu})P_{N,2}^*(z) = G(z)\bar{\mu}P_{N-1,2}(0) + G(z)\bar{\mu}P_{N,2}(0) - \bar{\mu}P_{N,2}(0). \quad (33)$$

将(23)–(33)式左右两端相加化简后可得如下一个重要等式

$$\begin{aligned} & \sum_{i=0}^N P_{i,0}^*(z) + \sum_{i=a}^{b-1} P_{i,1}^*(z) + \sum_{i=0}^N P_{i,2}^*(z) \\ &= \frac{G(z)-1}{z-1} \left[\sum_{i=0}^N P_{i,0}(0) + \sum_{i=a}^{b-1} P_{i,1}(0) + \sum_{i=0}^N P_{i,2}(0) \right]. \end{aligned}$$

当 $z \rightarrow 1^-$ 时对上式两端取极限, 并注意到 $P_{i,j}^*(1) = P_{i,j}$ ($j=0, 1, 2$), 由 L-Hospital 法则可得

$$\sum_{i=0}^N P_{i,0} + \sum_{i=a}^{b-1} P_{i,1} + \sum_{i=0}^N P_{i,2} = \frac{1}{\lambda} \left[\sum_{i=0}^N P_{i,0}(0) + \sum_{i=a}^{b-1} P_{i,1}(0) + \sum_{i=0}^N P_{i,2}(0) \right].$$

进一步, 结合归一化条件有

$$\lambda = \sum_{i=0}^N P_{i,0}(0) + \sum_{i=a}^{b-1} P_{i,1}(0) + \sum_{i=0}^N P_{i,2}(0). \quad (34)$$

令 $P_{i,0}^-$ 表示稳态下顾客到达前夕服务员处于休假状态且缓冲空间中顾客人数为 i 的概率, $P_{i,1}^-$ 表示稳态下顾客到达前夕服务员处于繁忙状态且正在服务的顾客人数为 i ($a \leq i < b$) 的概率, $P_{i,2}^-$ 表示稳态下顾客到达前夕服务员处于繁忙状态, 当前正在服务的顾客人数已达批量上限且缓冲空间中等待顾客人数为 i 的概率. 由 Bayes 公式及 (34) 式可知

$$\begin{aligned} P_{i,0}^- &= \lim_{t \rightarrow \infty} \text{Prob}\{N_q(t) = i, Y(t) = 0 \mid U(t) = 0\} \\ &= \lim_{t \rightarrow \infty} \frac{\text{Prob}\{N_q(t) = i, Y(t) = 0, U(t) = 0\}}{\text{Prob}\{U(t) = 0\}} \\ &= \frac{P_{i,0}(0)}{\sum_{i=0}^N P_{i,0}(0) + \sum_{i=a}^{b-1} P_{i,1}(0) + \sum_{i=0}^N P_{i,2}(0)} = \frac{P_{i,0}(0)}{\lambda}, \quad i = 0, \dots, N. \end{aligned} \quad (35)$$

同理可证

$$P_{i,1}^- = \frac{P_{i,1}(0)}{\lambda}, \quad i = a, \dots, b-1, \quad (36)$$

$$P_{i,2}^- = \frac{P_{i,2}(0)}{\lambda}, \quad i = 0, \dots, N. \quad (37)$$

利用 (35)–(37) 式, 并在 (25)–(33) 式中分别置变量 $z = 1$, 则稳态概率 $P_{i,2}$ ($i = 0, \dots, N$), $P_{i,0}$ ($i = a, \dots, N$), $P_{i,1}$ ($i = a, \dots, b-1$) 和顾客到达前夕概率之间的迭代关系可由 (38)–(46) 式给出

$$P_{N,2} = \frac{\lambda\bar{\mu}}{\mu} P_{N-1,2}^-, \quad (38)$$

$$P_{i,2} = \frac{\lambda\bar{\mu}}{\mu} (P_{i-1,2}^- - P_{i,2}^-), \quad i = N-b+1, \dots, N-1, \quad (39)$$

$$P_{N,0} = \frac{\lambda\bar{\theta}}{\theta} P_{N-1,0}^-, \quad (40)$$

$$P_{i,0} = \frac{\lambda\bar{\theta}}{\theta} (P_{i-1,0}^- - P_{i,0}^-), \quad i = a, \dots, N-1, \quad (41)$$

$$P_{N-b,2} = \frac{\theta}{\mu} P_{N,0} + P_{N,2} + \frac{\lambda\theta}{\mu} P_{N-1,0}^- + \lambda P_{N-1,2}^- + \frac{\lambda\bar{\mu}}{\mu} P_{N-b-1,2}^- - \frac{\lambda\bar{\mu}}{\mu} P_{N-b,2}^-, \quad (42)$$

$$\begin{aligned} P_{i,2} &= \frac{\theta}{\mu} P_{i+b,0} + P_{i+b,2} + \frac{\lambda\theta}{\mu} P_{i+b-1,0}^- + \lambda P_{i+b-1,2}^- + \frac{\lambda\bar{\mu}}{\mu} P_{i-1,2}^- \\ &\quad - \frac{\lambda\bar{\mu}}{\mu} P_{i,2}^- - \frac{\lambda\theta}{\mu} P_{i+b,0}^- - \lambda P_{i+b,2}^-, \quad i = 1, \dots, N-b-1, \end{aligned} \quad (43)$$

$$P_{0,2} = \frac{\theta}{\mu} P_{b,0} + P_{b,2} + \frac{\lambda\bar{\mu}}{\mu} P_{b-1,1}^- + \frac{\lambda\theta}{\mu} P_{b-1,0}^- + \lambda P_{b-1,2}^- - \frac{\lambda\bar{\mu}}{\mu} P_{0,2}^- - \frac{\lambda\theta}{\mu} P_{b,0}^- - \lambda P_{b,2}^-, \quad (44)$$

$$\begin{aligned} P_{i,1} &= P_{i,2} + \frac{\theta}{\mu} P_{i,0} + \frac{\lambda\theta}{\mu} P_{i-1,0}^- + \frac{\lambda\bar{\mu}}{\mu} P_{i-1,1}^- + \lambda P_{i-1,2}^- \\ &\quad - \frac{\lambda\bar{\mu}}{\mu} P_{i,1}^- - \lambda P_{i,2}^- - \frac{\lambda\theta}{\mu} P_{i,0}^-, \quad i = a+1, \dots, b-1, \end{aligned} \quad (45)$$

$$P_{a,1} = P_{a,2} + \frac{\theta}{\mu} P_{a,0} + \frac{\lambda\theta}{\mu} P_{a-1,0}^- + \lambda P_{a-1,2}^- - \frac{\lambda\bar{\mu}}{\mu} P_{a,1}^- - \lambda P_{a,2}^- - \frac{\lambda\theta}{\mu} P_{a,0}^- \quad (46)$$

稳态概率 $P_{i,0}$ ($i = 1, \dots, a-1$) 可由 (24) 式关于变量 z 在 $z = 1$ 处求导后得到

$$P_{i,0} = P_{i-1,0}^- + \mu P_{i,2}^{*(1)}(1) + \mu P_{i-1,2}^-, \quad i = 1, \dots, a-1, \quad (47)$$

这里 $P_{i,2}^{*(1)}(1)$ 表示 Z 变换 $P_{i,2}^*(z)$ 在 $z = 1$ 处的一阶导数, $P_{i,2}^{*(1)}(1)$ ($i = 1, \dots, N$) 的具体数值可由递推公式 (48)–(51) 获得

$$P_{N,2}^{*(1)}(1) = \frac{1}{\mu} (\bar{\mu} P_{N-1,2}^- + \bar{\mu} P_{N,2}^- - P_{N,2}), \quad (48)$$

$$P_{i,2}^{*(1)}(1) = \frac{1}{\mu} (\bar{\mu} P_{i-1,2}^- - P_{i,2}), \quad i = N-b+1, \dots, N-1, \quad (49)$$

$$P_{N-b,2}^{*(1)}(1) = \frac{1}{\mu} (\theta P_{N,0}^{*(1)}(1) + \mu P_{N,2}^{*(1)}(1) + \theta P_{N-1,0}^- + \theta P_{N,0}^- + \mu P_{N-1,2}^- + \mu P_{N,2}^- + \bar{\mu} P_{N-b-1,2}^- - P_{N-b,2}), \quad (50)$$

$$P_{i,2}^{*(1)}(1) = \frac{1}{\mu} (\theta P_{i+b,0}^{*(1)}(1) + \mu P_{i+b,2}^{*(1)}(1) + \theta P_{i+b-1,0}^- + \mu P_{i+b-1,2}^- + \bar{\mu} P_{i-1,2}^- - P_{i,2}), \quad i = N-b-1, \dots, 1. \quad (51)$$

类似 $P_{i,2}^{*(1)}(1)$ 的定义, (50), (51) 式中的 $P_{N,0}^{*(1)}(1)$ 和 $P_{i+b,0}^{*(1)}(1)$ ($i = N-b-1, \dots, 1$) 分别表示 Z 变换 $P_{N,0}^*(z)$ 和 $P_{i+b,0}^*(z)$ ($i = N-b-1, \dots, 1$) 在 $z = 1$ 处的一阶导数, 它们可由 (26) 和 (25) 式关于变量 z 在 $z = 1$ 处求导后得到

$$P_{N,0}^{*(1)}(1) = \frac{1}{\theta} (\bar{\theta} P_{N-1,0}^- + \bar{\theta} P_{N,0}^- - P_{N,0}), \quad (52)$$

$$P_{i,0}^{*(1)}(1) = \frac{1}{\theta} (\bar{\theta} P_{i-1,0}^- - P_{i,0}), \quad i = a, \dots, N-1. \quad (53)$$

最后稳态概率 $P_{0,0}$ 可由归一化条件依据如下公式获得

$$P_{0,0} = 1 - \sum_{i=1}^N P_{i,0} - \sum_{i=a}^{b-1} P_{i,1} - \sum_{i=0}^N P_{i,2}. \quad (54)$$

注 2 稳态概率 $P_{0,0}$ 也可以通过 (23) 式关于变量 z 在 $z = 1$ 处求导后得到. 在数值实验过程中我们发现两种方法求得的数值结果完全一致, 这从一个侧面说明了本文计算方法理论分析部分的正确性.

3.2 系统外部观察时刻与任意时刻队长分布的迭代关系

由于本文所研究的排队系统不遵循空竭服务原则, 且相邻顾客到达间隔时间服从一般离散分布, 其卷积运算难于实施, 因此给出顾客在缓冲空间中平均等待时间的解析表达式是相当困难的一件事. 研究系统外部观察时刻队长分布的主要目的是想利用 Little 公式, 从数值计算的途径来给出顾客在缓冲空间中的平均等待时间.

令 $P_{i,0}^{\text{out}}$ 表示稳态下系统外部观察时刻服务员处于休假状态且缓冲空间中顾客人数为 i 的概率, $P_{i,1}^{\text{out}}$ 表示稳态下系统外部观察时刻服务员处于繁忙状态且正在服务的顾客人数为 i ($a \leq i < b$) 的概率, $P_{i,2}^{\text{out}}$ 表示稳态下系统外部观察时刻服务员处于繁忙状态, 当前正在服务的顾客人数已达批量上限且缓冲空间中等待顾客人数为 i 的概率. 参考图 1, 并分析队长的演化过程, 将外部观察时刻与任意时刻即时隙分点处的状态联系起来, 容易建立起如下关系式

$$P_{i,2} = \bar{\mu}P_{i,2}^{\text{out}}, \quad i = N - b + 1, \dots, N, \quad (55)$$

$$P_{i,0} = \bar{\theta}P_{i,0}^{\text{out}}, \quad i = a, \dots, N, \quad (56)$$

$$P_{i,2} = \bar{\mu}P_{i,2}^{\text{out}} + \mu P_{i+b,2}^{\text{out}} + \theta P_{i+b,0}^{\text{out}}, \quad i = 0, \dots, N - b, \quad (57)$$

$$P_{i,0} = P_{i,0}^{\text{out}} + \mu P_{i,2}^{\text{out}}, \quad i = 1, \dots, a - 1, \quad (58)$$

$$P_{i,1} = \bar{\mu}P_{i,1}^{\text{out}} + \theta P_{i,0}^{\text{out}} + \mu P_{i,2}^{\text{out}}, \quad i = a, \dots, b - 1, \quad (59)$$

$$P_{0,0} = P_{0,0}^{\text{out}} + \mu \sum_{i=a}^{b-1} P_{i,1}^{\text{out}} + \mu P_{0,2}^{\text{out}}. \quad (60)$$

进一步由式 (55)–(60) 有

$$P_{i,2}^{\text{out}} = \frac{1}{\bar{\mu}}P_{i,2}, \quad i = N - b + 1, \dots, N, \quad (61)$$

$$P_{i,0}^{\text{out}} = \frac{1}{\bar{\theta}}P_{i,0}, \quad i = a, \dots, N, \quad (62)$$

$$P_{i,2}^{\text{out}} = \frac{1}{\bar{\mu}}P_{i,2} - \frac{\mu}{\bar{\mu}}P_{i+b,2}^{\text{out}} - \frac{\theta}{\bar{\mu}}P_{i+b,0}^{\text{out}}, \quad i = 0, \dots, N - b, \quad (63)$$

$$P_{i,0}^{\text{out}} = P_{i,0} - \mu P_{i,2}^{\text{out}}, \quad i = 1, \dots, a - 1, \quad (64)$$

$$P_{i,1}^{\text{out}} = \frac{1}{\bar{\mu}}P_{i,1} - \frac{\theta}{\bar{\mu}}P_{i,0}^{\text{out}} - \frac{\mu}{\bar{\mu}}P_{i,2}^{\text{out}}, \quad i = a, \dots, b - 1, \quad (65)$$

$$P_{0,0}^{\text{out}} = P_{0,0} - \mu \sum_{i=a}^{b-1} P_{i,1}^{\text{out}} - \mu P_{0,2}^{\text{out}}. \quad (66)$$

从 (38)–(66) 式不难发现只要获得顾客到达前夕系统稳态队长分布 $P_{i,j}^-$, 便可以通过上述迭代计算程序得到稳态下任意时刻和系统外部观察时刻的队长分布. 因此在 3.3 节中我们将进一步讨论概率 $P_{i,j}^-$ 的求解.

3.3 模型的嵌入 Markov 链与转移概率矩阵

用 T_n ($n = 1, 2, \dots$) 表示第 n 个顾客的到达时刻, 从而 $T_{n+1} - T_n$ ($n = 0, 1, 2, \dots, T_0 = 0$) 表示第 n 个与第 $n+1$ 个顾客之间的到达间隔时间. 记 $L_q(T_n^-)$ 表示第 n 个顾客到达前夕缓冲空间中排队等候的顾客数, $L_s(T_n^-)$ 表示第 n 个顾客到达前夕正在接受服务的顾客个数, $L_n^- = L_q(T_n^-) + L_s(T_n^-)$ 表示第 n 个顾客到达前夕系统中的顾客个数. 进一

步, 假设

$$J_n = \begin{cases} 0, & \text{第 } n \text{ 个顾客到达前夕系统处于休假状态;} \\ 1, & \text{第 } n \text{ 个顾客达到前夕服务台繁忙且此时} \\ & \text{正在服务的顾客人数未达到批量上限 } b; \\ 2, & \text{第 } n \text{ 个顾客达到前夕服务台繁忙且此时} \\ & \text{正在服务的顾客人数已达批量上限 } b. \end{cases}$$

由于假期中顾客只能在缓冲空间中等待, 所以当第 n 个顾客在假期中到达时 $L_n^- = L_q(T_n^-)$. 若第 n 个顾客到达时服务台处于繁忙状态且正在服务的顾客人数未达到批量上限, 这时第 n 个顾客观察到的系统缓冲空间中顾客个数 $L_q(T_n^-) \equiv 0$, 所以此时 $L_n^- = L_s(T_n^-)$. 最后, 若第 n 个顾客到达时服务台处于繁忙状态且正在服务的顾客人数已达批量上限, 这时原本 $L_n^- = L_q(T_n^-) + L_s(T_n^-)$, 但考虑到 $L_s(T_n^-) \equiv b$, 为减少系统状态描述的维数, 我们只记 $L_n^- = L_q(T_n^-)$, 这样的操作并不会引起系统状态个数的减少反而会简化我们的分析过程. 根据 Markov 过程的一般理论, 易知 $\{(L_n^-, J_n), n \geq 1\}$ 是该排队系统队长过程的一条嵌入 Markov 链, 有状态空间

$$\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3,$$

其中 $\Omega_1 = \{(i, 0) \mid i = 0, 1, \dots, N\}$, $\Omega_2 = \{(i, 1) \mid i = a, a+1, \dots, b-1\}$, $\Omega_3 = \{(i, 2) \mid i = 0, 1, \dots, N\}$. 将状态空间按照 J_n 水平的递增序排列, 可知嵌入 Markov 链的状态转移概率矩阵具有如下的分块结构

$$\mathbf{P} = \begin{matrix} & \begin{matrix} (0,0) & \cdots & (N,0) & (a,1) & \cdots & (b-1,1) & (0,2) & \cdots & (N,2) \end{matrix} \\ \begin{matrix} (0,0) \\ \vdots \\ (N,0) \\ (a,1) \\ \vdots \\ (b-1,1) \\ (0,2) \\ \vdots \\ (N,2) \end{matrix} & \begin{pmatrix} & & & \vdots & & & \vdots & & \\ & \mathbf{P}_{1,1} & & \vdots & \mathbf{P}_{1,2} & & \vdots & \mathbf{P}_{1,3} & \\ & & & \vdots & & & \vdots & & \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots \\ & \mathbf{P}_{2,1} & & \vdots & \mathbf{P}_{2,2} & & \vdots & \mathbf{P}_{2,3} & \\ & & & \vdots & & & \vdots & & \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots \\ & \mathbf{P}_{3,1} & & \vdots & \mathbf{P}_{3,2} & & \vdots & \mathbf{P}_{3,3} & \\ & & & \vdots & & & \vdots & & \\ & & & \vdots & & & \vdots & & \end{pmatrix}, \end{matrix}$$

为给出矩阵 \mathbf{P} 中元素 $p_{(i,m),(j,r)} = \text{Prob}\{L_{n+1}^- = j, J_{n+1} = r \mid L_n^- = i, J_n = m\}$ 的具体表达形式我们引入以下几个记号:

$$\alpha = \sum_{k=1}^{\infty} g_k (1-\theta)^k, \quad \gamma_l = \sum_{k=\max(l,1)}^{\infty} g_k \binom{k}{l} \mu^l (1-\mu)^{k-l}, \quad l = 0, 1, \dots,$$

$$\beta_l = \sum_{k=1+l}^{\infty} g_k \sum_{m=1}^{k-l} \theta(1-\theta)^m \binom{k-m}{l} \mu^l (1-\mu)^{k-m-l}, \quad l=0, 1, \dots$$

下面我们将利用上述记号对分块矩阵 \mathbf{P}_{k_1, k_2} ($k_1 = 1, 2, 3$, $k_2 = 1, 2, 3$) 中的元素逐一进行分析.

I 分块矩阵 $\mathbf{P}_{1,1}$ 中元素的分析.

- (1) 当 $i = 0, \dots, a-2$ 且 $j = i+1$ 时 $p_{(i,0),(j,0)} = 1$;
- (2) 当 $i = a-1, \dots, N-1$ 且 $j = i+1$ 时 $p_{(i,0),(j,0)} = \text{Prob}\{V > T_{n+1} - T_n\} = \alpha$;
- (3) 当 $i = a-1, \dots, N-1$ 且 i 除以批量服务上限 b 的余数在集合 $\{a-1, a, \dots, b-1\}$ 中取值时

$$p_{(i,0),(0,0)} = 1 - \alpha - \sum_{j=a}^{b-1} p_{(i,0),(j,1)} - \sum_{j=0}^N p_{(i,0),(j,2)};$$

- (4) 当 $i = b, \dots, N-1$, $j = 1, \dots, a-1$ 且 $\frac{i+1-j}{b}$ 为一正整数时

$$p_{(i,0),(j,0)} = 1 - \alpha - \sum_{r=a}^{b-1} p_{(i,0),(r,1)} - \sum_{r=0}^N p_{(i,0),(r,2)};$$

- (5) 当 $j = 0, \dots, N$ 时 $p_{(N,0),(j,0)} = p_{(N-1,0),(j,0)}$;

- (6) 其余情形下 $p_{(i,0),(j,0)} = 0$.

II 分块矩阵 $\mathbf{P}_{1,2}$ 中元素的分析.

- (1) 当 $i = a-1, \dots, N-1$, $j = a, \dots, b-1$ 且 i 除以批量服务上限 b 的余数加 1 恰为 j 时

$$\begin{aligned} p_{(i,0),(j,1)} &= \text{Prob}\left\{V + \sum_{l=1}^{\lfloor \frac{i}{b} \rfloor} S_l \leq T_{n+1} - T_n < V + \sum_{l=1}^{\lfloor \frac{i}{b} \rfloor + 1} S_l\right\} \\ &= \sum_{k=\lfloor \frac{i}{b} \rfloor + 1}^{\infty} g_k \text{Prob}\left\{V + \sum_{l=1}^{\lfloor \frac{i}{b} \rfloor} S_l \leq k < V + \sum_{l=1}^{\lfloor \frac{i}{b} \rfloor + 1} S_l\right\} \\ &= \sum_{k=\lfloor \frac{i}{b} \rfloor + 1}^{\infty} g_k \sum_{m=1}^{k-\lfloor \frac{i}{b} \rfloor} \theta(1-\theta)^{m-1} \text{Prob}\left\{\sum_{l=1}^{\lfloor \frac{i}{b} \rfloor} S_l \leq k-m < \sum_{l=1}^{\lfloor \frac{i}{b} \rfloor + 1} S_l\right\} \\ &= \sum_{k=\lfloor \frac{i}{b} \rfloor + 1}^{\infty} g_k \sum_{m=1}^{k-\lfloor \frac{i}{b} \rfloor} \theta(1-\theta)^{m-1} \binom{k-m}{\lfloor \frac{i}{b} \rfloor} \mu^{\lfloor \frac{i}{b} \rfloor} (1-\mu)^{k-m-\lfloor \frac{i}{b} \rfloor} \\ &= \beta_{\lfloor \frac{i}{b} \rfloor}; \end{aligned}$$

这里 $[x]$ 表示不大于 x 的最大整数, 约定当 $n < m$ 时 $\sum_{l=m}^n = 0$.

- (2) 当 $j = a, \dots, b-1$ 时 $p_{(N,0),(j,1)} = p_{(N-1,0),(j,1)}$;

- (3) 其余情形下 $p_{(i,0),(j,1)} = 0$.

III 分块矩阵 $\mathbf{P}_{1,3}$ 中元素的分析.

(1) 当 $i = b - 1, \dots, N - 1$, $j = 0, \dots, N - b$ 且 $\frac{i+1-j}{b}$ 为一正整数时

$$\begin{aligned}
 p_{(i,0),(j,2)} &= \text{Prob} \left\{ V + \sum_{l=1}^{\frac{i+1-j}{b}-1} S_l \leq T_{n+1} - T_n < V + \sum_{l=1}^{\frac{i+1-j}{b}} S_l \right\} \\
 &= \sum_{k=\frac{i+1-j}{b}}^{\infty} g_k \text{Prob} \left\{ V + \sum_{l=1}^{\frac{i+1-j}{b}-1} S_l \leq k < V + \sum_{l=1}^{\frac{i+1-j}{b}} S_l \right\} \\
 &= \sum_{k=\frac{i+1-j}{b}}^{\infty} g_k \sum_{m=1}^{k-\frac{i+1-j}{b}+1} \theta(1-\theta)^{m-1} \text{Prob} \left\{ \sum_{l=1}^{\frac{i+1-j}{b}-1} S_l \leq k-m < \sum_{l=1}^{\frac{i+1-j}{b}} S_l \right\} \\
 &= \sum_{k=\frac{i+1-j}{b}}^{\infty} g_k \sum_{m=1}^{k-\frac{i+1-j}{b}+1} \theta(1-\theta)^{m-1} \binom{k-m}{\frac{i+1-j}{b}-1} \mu^{\frac{i+1-j}{b}-1} (1-\mu)^{k-m-\frac{i+1-j}{b}+1} \\
 &= \beta_{\frac{i+1-j}{b}-1};
 \end{aligned}$$

(2) 当 $j = 0, 1, \dots, N$ 时 $p_{(N,0),(j,2)} = p_{(N-1,0),(j,2)}$;

(3) 其余情形下 $p_{(i,0),(j,2)} = 0$.

IV 分块矩阵 $\mathbf{P}_{2,1}$ 中元素的分析.

(1) 当 $i = a, \dots, b - 1$ 时 $p_{(i,1),(0,0)} = 1 - \gamma_0$;

(2) 其余情形下 $p_{(i,1),(j,0)} = 0$.

V 分块矩阵 $\mathbf{P}_{2,2}$ 中元素的分析.

(1) 当 $i = a, \dots, b - 2$, $j = i + 1$ 时 $p_{(i,1),(j,1)} = \text{Prob}\{S_1 > T_{n+1} - T_n\} = \gamma_0$;

(2) 其余情形下 $p_{(i,1),(j,1)} = 0$.

VI 分块矩阵 $\mathbf{P}_{2,3}$ 中元素的分析.

(1) 当 $i = b - 1$, $j = 0$ 时 $p_{(b-1,1),(0,2)} = \text{Prob}\{S_1 > T_{n+1} - T_n\} = \gamma_0$;

(2) 其余情形下 $p_{(i,1),(j,2)} = 0$.

VII 分块矩阵 $\mathbf{P}_{3,1}$ 中元素的分析.

(1) 当 $i = 0, \dots, N - 1$, $j = 1, \dots, a - 1$ 且 $\frac{i+1-j}{b}$ 为一非负整数时

$$p_{(i,2),(j,0)} = 1 - \sum_{r=a}^{b-1} p_{(i,2),(r,1)} - \sum_{r=0}^N p_{(i,2),(r,2)};$$

(2) 当 $i = a - 1, \dots, N - 1$ 且 i 除以 b 的余数在集合 $\{a - 1, a, \dots, b - 1\}$ 中取值时

$$p_{(i,2),(0,0)} = 1 - \sum_{r=a}^{b-1} p_{(i,2),(r,1)} - \sum_{r=0}^N p_{(i,2),(r,2)};$$

(3) 当 $j = 0, 1, \dots, N$ 时 $p_{(N,2),(j,0)} = p_{(N-1,2),(j,0)}$;

(4) 其余情形 $p_{(i,2),(j,0)} = 0$.

VIII 分块矩阵 $\mathbf{P}_{3,2}$ 中元素的分析.

(1) 当 $i = a - 1, \dots, N - 1, j = a, \dots, b - 1$ 且 i 除以 b 的余数加 1 恰为 j 时

$$\begin{aligned} p_{(i,2),(j,1)} &= \text{Prob} \left\{ \sum_{l=1}^{[\frac{i}{b}]+1} S_l \leq T_{n+1} - T_n < \sum_{l=1}^{[\frac{i}{b}]+2} S_l \right\} \\ &= \sum_{k=[\frac{i}{b}]+1}^{\infty} g_k \text{Prob} \left\{ \sum_{l=1}^{[\frac{i}{b}]+1} S_l \leq k < \sum_{l=1}^{[\frac{i}{b}]+2} S_l \right\} \\ &= \sum_{k=[\frac{i}{b}]+1}^{\infty} g_k \binom{k}{[\frac{i}{b}]+1} \mu^{[\frac{i}{b}]+1} (1-\mu)^{k-[\frac{i}{b}]-1} = \gamma_{[\frac{i}{b}]+1}; \end{aligned}$$

(2) 当 $j = a, \dots, b - 1$ 时 $p_{(N,2),(j,1)} = p_{(N-1,2),(j,1)}$;

(3) 其余情形 $p_{(i,2),(j,1)} = 0$.

IX 分块矩阵 $\mathbf{P}_{3,3}$ 中元素的分析.

(1) 当 $i = 0, \dots, N - 1, j = i + 1$ 时

$$p_{(i,2),(j,2)} = \text{Prob}\{S_1 > T_{n+1} - T_n\} = \gamma_0;$$

(2) 当 $i = b - 1, \dots, N - 1, j = 0, \dots, N - b$ 且 $\frac{i+1-j}{b}$ 为正整数时

$$\begin{aligned} p_{(i,2),(j,2)} &= \text{Prob} \left\{ \sum_{l=1}^{\frac{i+1-j}{b}} S_l \leq T_{n+1} - T_n < \sum_{l=1}^{\frac{i+1-j}{b}+1} S_l \right\} \\ &= \sum_{k=\frac{i+1-j}{b}}^{\infty} g_k \text{Prob} \left\{ \sum_{l=1}^{\frac{i+1-j}{b}} S_l \leq k < \sum_{l=1}^{\frac{i+1-j}{b}+1} S_l \right\} \\ &= \sum_{k=\frac{i+1-j}{b}}^{\infty} g_k \binom{k}{\frac{i+1-j}{b}} \mu^{\frac{i+1-j}{b}} (1-\mu)^{k-\frac{i+1-j}{b}} = \gamma_{\frac{i+1-j}{b}}; \end{aligned}$$

(3) $j = 0, \dots, N$ 时 $p_{(N,2),(j,2)} = p_{(N-1,2),(j,2)}$;

(4) 其余情形 $p_{(i,2),(j,2)} = 0$.

令 $\mathcal{P} = \{P_{0,0}^-, \dots, P_{N,0}^-, P_{a,1}^-, \dots, P_{b-1,1}^-, P_{0,2}^-, \dots, P_{N,2}^-\}$, \mathbf{e} 表示一个 $2(N+1) + (b-a)$ 维的全 1 列向量. 运用 MATLAB 数值计算软件编程求解下述带有归一化条件的线性方程组:

$$\begin{cases} \mathcal{P}\mathbf{P} = \mathcal{P}, \\ \mathcal{P}\mathbf{e} = 1, \end{cases}$$

便可求得顾客到达前夕系统的稳态队长分布. 上述线性方程组计算程序的核心在于要按照转移概率矩阵的特殊结构将矩阵元素准确定位于 \mathbf{P} 中, 这一点可以通过 MATLAB 中的元胞数组 (cell) 方便的实现.

4 性能指标的数值结果

4.1 PH 到达间隔时间下的三种队长分布

在离散时间排队系统中, 不同时刻的队长分布是诸多排队指标中最为重要的一类性能指标. 在获知系统不同时刻队长分布后, 我们可以直接利用它来获得平均队长, 顾客阻塞概率等工程实践中较为关注的性能测度. 为展示本文队长分布迭代算法的具体应用, 我们利用 MATLAB 软件编写数值计算程序, 给出到达间隔时间为 PH 分布情形下的 $PH/Geom^{(4,8)}/1/23$ 排队系统三种时刻的队长分布 (见表 1). 实验中数值参数的设置如下: $\mu = 0.07, \theta = 0.02$ 且到达过程是具有不可约表示 $(\mathbf{T}, \boldsymbol{\omega})$ 的 3 阶离散 PH 更新过程, 其中

$$\mathbf{T} = \begin{pmatrix} 0.15 & 0.3 & 0.25 \\ 0.4 & 0.3 & 0.1 \\ 0.2 & 0.1 & 0.4 \end{pmatrix}, \quad \boldsymbol{\omega} = \{0.3, 0.1, 0.6\},$$

文中所有数值实验的计算结果均保留六位小数.

表 1 $PH/Geom^{(4,8)}/1/23$ 排队系统三种时刻的稳态队长分布

i	$P_{i,0}$	$P_{i,0}^-$	$P_{i,0}^{out}$	$P_{i,1}$	$P_{i,1}^-$	$P_{i,1}^{out}$	$P_{i,2}$	$P_{i,2}^-$	$P_{i,2}^{out}$
0	0.003922	0.007387	0.001882	-	-	-	0.016896	0.016665	0.016146
1	0.009905	0.010086	0.009160	-	-	-	0.019307	0.019094	0.018636
2	0.012940	0.013099	0.012107	-	-	-	0.021380	0.021185	0.020802
3	0.016234	0.016374	0.015329	-	-	-	0.023091	0.022914	0.022614
4	0.015229	0.015249	0.015540	0.004490	0.004125	0.003350	0.024439	0.024279	0.024062
5	0.014182	0.014201	0.014472	0.008123	0.007787	0.007112	0.025435	0.025291	0.025156
6	0.013208	0.013225	0.013477	0.011315	0.011006	0.010426	0.026102	0.025972	0.025914
7	0.012300	0.012316	0.012551	0.014783	0.013950	0.013970	0.028537	0.027837	0.028022
8	0.011455	0.011470	0.011688	-	-	-	0.029479	0.029391	0.029050
9	0.010667	0.010681	0.010885	-	-	-	0.030159	0.030090	0.029966
10	0.009934	0.009947	0.010137	-	-	-	0.030181	0.030127	0.030171
11	0.009252	0.009264	0.009440	-	-	-	0.029704	0.029662	0.029833
12	0.008616	0.008627	0.008792	-	-	-	0.028858	0.028825	0.029089
13	0.008024	0.008034	0.008187	-	-	-	0.027744	0.027719	0.028050
14	0.007472	0.007482	0.007625	-	-	-	0.026447	0.026428	0.026803
15	0.006959	0.006968	0.007101	-	-	-	0.041202	0.040394	0.037345
16	0.006480	0.006489	0.006613	-	-	-	0.035019	0.035110	0.036479
17	0.006035	0.006043	0.006158	-	-	-	0.030439	0.030518	0.031707
18	0.005620	0.005628	0.005735	-	-	-	0.026458	0.026526	0.027560
19	0.005234	0.005241	0.005341	-	-	-	0.022997	0.023057	0.023955
20	0.004874	0.004881	0.004974	-	-	-	0.019989	0.020041	0.020822
21	0.004539	0.004545	0.004632	-	-	-	0.017375	0.017420	0.018099
22	0.004227	0.004233	0.004314	-	-	-	0.015102	0.015141	0.015731
23	0.057281	0.057356	0.058450	-	-	-	0.100361	0.100621	0.104542

由表 1 中的数据不难发现, 当服务员繁忙时, 在稳态下三种时刻的队长分布中, 系统处于状态 $(N-b, 2)$ (即 $(15, 2)$) 的概率最大, 这和 3.1 节中对系统状态转移过程的概率分析是十分吻合的. 从 3.1 节的差分方程 (9) 中可以看到, 在服务员繁忙期间系统可以经由多个状态在一个单位时隙后转移至状态 $(N-b, u-1, 2)$, 正是基于这样的原因使得稳态下服务员繁忙时 $P_{N-b,2}$, $P_{N-b,2}^-$, $P_{N-b,2}^{out}$ 在同类概率中取值最大.

4.2 系统阻塞概率

由于本文讨论的排队系统缓冲空间容量有限, 因此当某位到达的顾客看到缓冲空间中顾客已满时他将离开系统. 我们将顾客看到缓冲空间已满的概率称为系统的阻塞概率, 记为 P_{loss} , 由 3.3 节的分析可知

$$P_{loss} = P_{N,0}^- + P_{N,2}^-.$$

下面我们将通过数值实验来考察在均值相同的定长 (Deterministic) 和几何 (Geometric) 输入流下, 系统阻塞概率随缓冲空间容量 N 和休假参数 θ 的变化. 图 2(a) 中实验参数的设置为: $\mu = 0.04$, $\theta = 0.02$, $a = 4$, $b = 8$, 其中, 几何分布的概率分布律为 $g_k = 0.2(1-0.2)^{k-1}$, $k \geq 1$, 定长分布的概率分布律为 $g_5 = 1$; 图 2(b) 中的数值实验只须固定缓冲容量 $N = 23$ 令 θ 发生变化, 其余参数取值与概率分布律的设置均与图 2(a) 中保持一致.

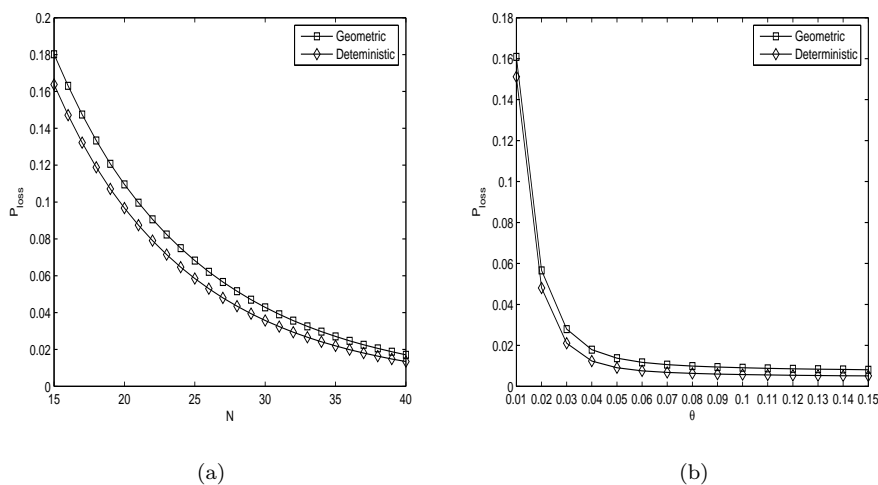


图 2. 两种到达间隔分布下系统阻塞概率随 N 和 θ 的变化

由图 2(a)–(b) 可以直观看出在到达间隔时间均值相同的情况下, 定长输入时的系统阻塞概率要比几何输入时的低, 这说明在随机服务系统的动态管理中应尽量让输入端保持一种持续稳定的状态, 以减小由于输入的不确定性所带来的系统拥塞. 图 2(a) 显示系统阻塞概率随缓冲空间容量的增大而减小, 但减小的幅度逐渐趋缓, 这说明仅仅采用单纯增加缓冲容量的办法来改善系统的拥塞状况并不可取; 图 2(b) 显示适当增大 θ

的取值即缩短服务员假期长度将有效降低系统的阻塞概率, 当 θ 的取值从 0.01 提高到 0.04(平均休假时间由 100 减小到 25 时) 系统的阻塞概率大致能下降 14% 左右, 但当 θ 的取值超过 0.05 后 θ 的变化对系统阻塞概率的影响不显著.

4.3 批量服务中途准入概率

当某位顾客到达系统, 发现服务员处于繁忙状态且当前正在接受服务的顾客数量大于等于 a 而小于 b 时, 则 he 可以直接加入当前的服务群体接受服务而不在缓冲空间中等待. 我们把这样一种事件发生的概率称为批量服务中途准入概率, 记为 $P_{\text{accessible}}$. 由 3.3 节的分析可知

$$P_{\text{accessible}} = \sum_{i=a}^{b-1} P_{i,1}^-.$$

图 3 研究了批量服务上限值 b 的变化对批量服务中途准入概率的影响. 在图 3 的数值实验中固定缓冲空间容量 $N = 42$, 批量服务上限值 b 的变化范围限制在 8 与 17 之间, 其余参数取值和概率分布律的设置均与图 2(a) 中相同. 图 3 显示在到达间隔时间均值相同的情况下, 定长输入时的批量服务中途准入概率要比几何输入时的高, 这恰好从一个侧面解释了图 2 中的实验结果, 正因为定长输入情形下的中途准入概率较几何输入时高所以导致其阻塞概率比几何输入时低.

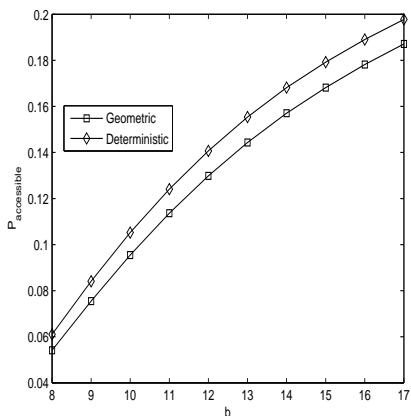


图 3. 批量上限值 b 对 $P_{\text{accessible}}$ 的影响

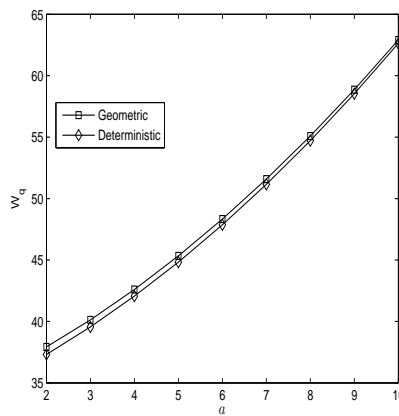


图 4. 服务开始阈值 a 对 W_q 的影响

4.4 顾客在缓冲空间中的平均等待时间

利用 3.2 节所提供的迭代公式可以获得系统外部观察时刻的队长分布, 利用该分布通过 Little 公式可以直接获得顾客在缓冲空间中的平均等待时间 W_q ,

$$W_q = \frac{\sum_{i=1}^N i(P_{i,0}^{\text{out}} + P_{i,2}^{\text{out}})}{\lambda(1 - P_{\text{loss}})},$$

其中 $\lambda(1 - P_{\text{loss}})$ 表示系统的有效到达率.

图 4 研究了服务开始阈值 a 的变化对顾客在缓冲空间中平均等待时间的影响. 图 4 中数值实验参数的设置如下: $N = 27$, $b = 14$, 阈值 a 的变化范围限制在 2 与 10 之间, 其余参数取值和概率分布律的设置均与图 2(a) 中一致. 由图 4 可知服务开始的阈值 a 越低顾客在缓冲空间中停留的时间也相应越短; 在到达间隔时间均值相同的情况下, 定长输入时顾客在缓冲空间中的平均等待时间比几何输入时的略小, 由于定长输入下的等待时间短所以缓冲空间中发生顾客溢出的机率也相对要小一些, 这也与图 2 中显示的结果是吻合的.

4.5 一种特殊情形下的数值算例

Samanta 等人在 [13] 中研究了离散时间 $GI/Geom/1/N$ 多重休假 EAS 排队系统, 这种排队系统实质上是本文所研究排队系统的一个特例. 我们注意到当把门限阈值 a 和批量服务上限值 b 都取为 1 时, 本文所研究的具有中途准入机制和多重休假的离散时间 $GI/Geom^{(a,b)}/1/N$ 早到排队系统便退化成了 Samanta 所研究的系统. 为进一步说明本文理论部分分析的正确性与计算程序编写的可靠性, 我们利用 [13] 表 2 中所罗列的系统参数进行了一个特例的数值实验. 值得注意的是在进行这个特例的数值实验时, 系统原本所呈现的某些状态会消失. 由于当 $a = 1$, $b = 1$ 时中途准入机制不复存在, 所以在进行这个特例的数值实验时是没有 $P_{i,1}$ 这种状态概率的, 仅留下了 $P_{i,0}$ 和 $P_{i,2}$ 两种状态概率. 而且更值得注意的是就剩下的这两种状态概率而言, 它们与 Samanta 等人在 [13] 中定义的状态概率还有一定的差别. 在 Samanta 等人定义的状态概率中 N_t 记录的是 t 时刻系统中的顾客数量 (包含正在接受服务的那一个在内), 而在我们的论文中 $N_q(t)$ 记录的是 t 时刻缓冲空间中的顾客数量, 所以本文中的状态概率 $P_{i,j}$, $P_{i,j}^-$, $P_{i,j}^{\text{out}}$ 与 [13] 中定义的 $Q_{i,j}$, $Q_{i,j}^-$, $Q_{i,j}^o$ 之间会有如下的数量关系成立:

$$\begin{cases} P_{i,0} = Q_{i,0}, & i = 0, 1, \dots, N, \\ P_{i,2} = Q_{i+1,1}, & i = 0, 1, \dots, N, \\ P_{i,0}^- = Q_{i,0}^-, & i = 0, 1, \dots, N, \\ P_{i,2}^- = Q_{i+1,1}^-, & i = 0, 1, \dots, N, \\ P_{i,0}^{\text{out}} = Q_{i,0}^o, & i = 0, 1, \dots, N, \\ P_{i,2}^{\text{out}} = Q_{i+1,1}^o, & i = 0, 1, \dots, N, \end{cases}$$

其中 $Q_{i,0}$, $Q_{i,0}^-$ 和 $Q_{i,0}^o$ 在文献 [13] 中分别表示稳态下任意时刻, 顾客到达前夕和系统外部观察时刻顾客数量为 i 且服务员正在休假的概率. 进一步, $Q_{i,1}$, $Q_{i,1}^-$ 和 $Q_{i,1}^o$ 在 [13] 中分别代表稳态下任意时刻, 顾客到达前夕和系统外部观察时刻顾客数量为 i 且服务员正处于繁忙状态的概率.

下面, 假设相邻两顾客的到达间隔时间分布为: $g_1 = 0.35$, $g_2 = 0.2$, $g_5 = 0.25$, $g_{11} = 0.15$, $g_{15} = 0.05$. 服务时间参数和休假参数分别为 $\mu = 0.24$, $\theta = 0.018$, 系统容量 $N = 170$, 阈值 a 和批量服务上限值 b 均设置为 1. 将原来用 Matlab 编写的数值计算

程序源代码做相应调整后, 计算结果如下表 2 所示. 由于实验所取缓冲空间容量较大 ($N = 170$), 所以在表 2 中我们只罗列了 i 取值为 1 到 9 的情形.

表 2 一种特殊情形下排队系统三种时刻的稳态队长分布

i	$Q_{i,0}^-$	$Q_{i,0}$	$Q_{i,0}^o$	$P_{i,0}^-$	$P_{i,0}$	$P_{i,0}^{\text{out}}$
0	0.007513	0.006736	0.005103	0.007513	0.006736	0.005103
1	0.006929	0.006923	0.006923	0.006929	0.006923	0.006923
2	0.006391	0.006385	0.006385	0.006391	0.006385	0.006385
3	0.005894	0.005889	0.005889	0.005894	0.005889	0.005889
4	0.005436	0.005432	0.005432	0.005436	0.005432	0.005432
5	0.005041	0.005010	0.005010	0.005041	0.005010	0.005010
6	0.004624	0.004620	0.004620	0.004624	0.004620	0.004620
7	0.004265	0.004261	0.004261	0.004265	0.004261	0.004261
8	0.003933	0.003930	0.003930	0.003933	0.003930	0.003930
9	0.003628	0.003625	0.003625	0.003628	0.003625	0.003625

i	$Q_{i,1}^-$	$Q_{i,1}$	$Q_{i,1}^o$	$P_{i,2}^-$	$P_{i,2}$	$P_{i,2}^{\text{out}}$
0	-	-	-	0.008301	0.008483	0.006805
1	0.008301	0.008483	0.006805	0.014976	0.015130	0.013796
2	0.014976	0.015130	0.013796	0.020266	0.020396	0.019354
3	0.020266	0.020396	0.019354	0.024382	0.024491	0.023696
4	0.024382	0.024491	0.023696	0.027505	0.027596	0.027009
5	0.027505	0.027596	0.027009	0.029792	0.029868	0.029456
6	0.029792	0.029868	0.029456	0.031378	0.031441	0.031174
7	0.031378	0.031441	0.031174	0.032380	0.032431	0.032285
8	0.032380	0.032431	0.032285	0.032896	0.032938	0.032892
9	0.032896	0.032938	0.032892	0.033014	0.033048	0.033083

注意到上面所提及的 $P_{i,j}$, $P_{i,j}^-$, $P_{i,j}^{\text{out}}$ 与 $Q_{i,j}$, $Q_{i,j}^-$, $Q_{i,j}^o$ 之间的关系, 由表 2 中数据我们发现利用本文第 3 节提出的迭代公式编写的稳态概率数值计算程序所求得的队长分布与 Samanta 在文献 [13] 表 2 中所报道的结果完全一致.

5 结论

以往大多数分析离散时间排队模型的文献仅考虑了顾客到达前夕的系统稳态队长分布, 对任意时刻和系统外部观察时刻的队长分布未能作出进一步的探讨. 在本文的研究中我们综合运用补充变量和嵌入 Markov 链方法考虑了三种不同时刻的队长分布, 并通过系统外部观察时刻的稳态队长分布轻松获得了顾客在缓冲空间中的等待时间均值, 避免了较为繁琐和不易实施的等待时间均值的解析讨论, 为工程实践中排队指标的具体计算提供了一种行之有效的方法, 也使得本文的研究不仅具有理论意义, 也具有潜在的应用价值.

另外, 本文在研究 EAS 系统时选取的状态记录时刻是时隙分点 t , 这是研究离散时间 EAS 排队系统所惯常选取的状态记录时刻. 而在研究离散时间 LAS-DA (晚到延迟进

入) 排队系统时惯常选取的状态记录时刻却是 t^- (见 [13–15]). 正是这样的一点细节之处的不同就会导致在利用补充变量方法推导两种系统的状态演化差分方程组时产生较大的差异. 受文章篇幅所限我们在该文中只讨论了 EAS 系统, 对 LAS-DA 系统的探讨完全可仿照本文的研究而展开.

参 考 文 献

- [1] Gold H, Tran-Gia P. Performance Analysis of a Batch Service Queue Arising Out of Manufacturing System Modelling. *Queueing System*, 1993, 14(3-4): 413–426
- [2] Chaudhry M L, Templeton J G C. A First Course in Bulk Queues. New York: Wiley, 1983
- [3] Vijaya Laxmi P, Gupta U C. On the Finite-buffer Bulk-service Queue with General Independent Arrivals: $GI/M^{(b)}/1/N$. *Operations Research Letters*, 1999, 25(5): 241–245
- [4] Sivasamy R. A Bulk Service Queue with Accessible and Non-accessible Batches. *Opsearch*, 1990, 27(1): 46–54
- [5] Samanta S K, Chaudhry M L, Gupta U C. Discrete-time $Geo^X/G^{(a,b)}/1/N$ Queues with Single and Multiple Vacations. *Mathematical and Computer Modelling*, 2007(1-2), 45: 93–108
- [6] Gupta U C, Sikdar K. The Finite-buffer $M/G/1$ Queue with General Bulk-service Rule and Single Vacation. *Performance Evaluation*, 2004, 57(2): 199–219
- [7] Arumuganathan R, Rathinasamy A. Analysis of a Bulk Queue with Multiple Vacations and Accessible Batch Service. *Proceedings of the Second National Conference on Mathematical and Computational Models*, 2003: 79–88
- [8] Vijaya Laxmi P, Goswami V, Yesuf O M. Finite Buffer Single Vacation Queue with Accessible and Non-accessible Batch Service. *International Journal of Information and Management Sciences*, 2010, 21(2): 125–142
- [9] Goswami V, Mohanty J R, Samanta S K. Discrete-time Bulk-service Queues with Accessible and Non-accessible Batches. *Applied Mathematics and Computation*, 2006, 182(1): 898–906
- [10] Hunter J J. Mathematical Techniques of Applied Probability (Vol.2): Discrete Time Models: Techniques and Applications. New York: Academic Press, 1983
- [11] 田乃硕, 徐秀丽, 马占友. 离散时间排队论. 北京: 科学出版社, 2008
(Tian N S, Xu X L, Ma Z Y. Discrete-time Queueing Theory. Beijing: Science Press, 2008)
- [12] Gupta U C, Goswami V. Performance Analysis of Finite Buffer Discrete-time Queue with Bulk Service. *Computers and Operations Research*, 2002, 29(10): 1331–1341
- [13] Samanta S K, Gupta U C, Sharma R K. Analysis of Finite Capacity Discrete-time $GI/Geo/1$ Queueing System with Multiple Vacations. *Journal of the Operational Research Society*, 2007, 58(3): 368–377
- [14] Chaudhry M L, Gupta U C. Performance Analysis of the Discrete-time $GI/Geom/1/N$ Queue. *Journal of Applied Probability*, 1996, 33(1): 239–255
- [15] Tian N S, Zhang Z G. The Discrete-time $GI/Geo/1$ Queue with Multiple Vacations. *Queueing System*, 2002, 40(3): 283–294

GI/Geom^(a,b)/1/N Early Arrival Queueing System
with Accessible Mechanism of Ongoing Service
and Multiple Vacations

YU MIAOMIAO

(School of Mathematics & Software Science, Sichuan Normal University, Chengdu 610068)

(School of science, Sichuan University of Science and Engineering, Zigong 643000)

TANG YINGHUI

(School of Mathematics & Software Science, Sichuan Normal University, Chengdu 610068)

(E-mail: tangyh@uestc.edu.cn)

FU YONGHONG

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract Applying the supplementary variable technique and embedded Markov chain, a discrete-time multiple vacations queueing system with accessible mechanism of ongoing bulk service is considered. Through an effective iterated algorithm, we obtain the queue length distributions at three kinds of time epochs. Furthermore, using the queue length distribution at different time epochs and the numerical method, we study the influence of system parameters on several performance measures such as loss probability, accessible probability of ongoing bulk service and the mean waiting time of customers in the buffer space.

Key words discrete-time queue; accessible mechanism of ongoing bulk service;
supplementary variable technique; embedded Markov chain

MR(2000) Subject Classification 60K25; 90B22

Chinese Library Classification O226