第 9 卷 第 1 期　　运 筹 与 管 理　　Vol. 9, No. 1
2000 年 3 月　　OPERATIONS RESEARCH AND MANAGEMENT SCIENCE　　Mar., 2000

一类批量服务排队系统的求解

周志中[1], 周亚平[2]

(1. 清华大学 经济管理学院, 北京 100084; 2. 中国科学技术大学 商学院, 安徽 合肥 230027)

摘　要: 给出一类特殊的批量服务排队系统——公交系统的稳态概率分布的求解过程, 并在此基础上给出这类服务系统平均队长的算法。文中给出具体例子说明计算过程。

关键词: 排队论; 随机服务系统; 马尔可夫链; 批量服务

中图分类号: O226　　**文献标识码**: A　　**文章编号**: 1007-3221(2000)01-0021-05

The Solution to a Kind of Bulk Service Queueing System

ZHOU Zhi-zhong[1], ZHOU Ya-ping[2]

(1. *Economic & Management School*, *Tsinghua Univ.*, *Beijing* 100084, *China*; 2. *Business School of USTC*, *Hefei* 230027, *China*)

Abstract: A kind of special bulk-service queueing system is discussed in the paper and the solution to the steady-state probability is given. On the basis of the solution, we discuss the method of computing the mean line length. Many examples are given to show the process of computation.

Key wofrds: queueing theory; random service system; Markov chain; bulk service

0　引言

对于批量服务系统, 许多学者做了研究。Leonard Kleinrock[1]给出了一种特殊的批量服务系统 $M/M^K/1$ 的平稳概率分布: $\pi_K=(1-\frac{1}{z_0})(\frac{1}{z_0})^k$, 其中 z_0 是满足 $\lambda z^{K+1}-(\lambda+\mu)\cdot z^K+\mu=0$ 以及 $|z_0|>1$ 的解。徐光辉[2]也对 $M/M^K/1$ 做了探讨, 并给出类似结果。但他们所讨论的批量服务系统和现实中一种特殊的批量服务系统——公交系统略有出入, 公交服务系统服务的产生不依赖于系统的状态, 打比方说, 一辆列车刚刚开过, 即使站台上空无一人, 下辆列车的到达间隔已经开始计时了。而 Leonard Kleinrock 和徐光辉所指的服务系统和传统的服务系统一

收稿日期: 1999-07-08
基金项目: 中国科技大学青年基金资助项目(KA2612)。
作者简介: 周志中(1975-), 男, 海南海口人, 清华大学经济管理学院硕士研究生。

致，即：系统非空时服务间隔才开始计时。这两者的区别在计算机仿真时尤为明显。本文主要讨论一种特殊的批量服务排队系统——公交系统的求解。J. Medhi[3]给出了这类服务系统概率母函数的形式，但未能进一步探讨母函数中未知数的求法。J. W. Cohen[4]在其经典著作《The Single Server Queue》中给出批量服务系统各项指标的精确数学表达形式，然而 J. W. Cohen 在求解中使用了大量艰涩难懂的数学工具，因此所得结果也仅有理论上的意义。本文给出了对这类系统最重要的参数——稳态概率分布的一般求法，所用的数学工具也较为简单，有较强的操作性。

对于期望队长，华兴[5]给出了 $G/G/1$ 系统期望队长的一种求法，避开了求解 $\sum_{k=0}^{\infty} k \cdot \pi_k$ 的繁琐过程，在 π_k 的精确表达式无法得到的情况下，这种方法尤为有用。本文参考了华兴的思路，给出求解期望队长的一般解法。本文的批量服务系统比较特殊，在求出 $\pi_0, \cdots, \pi_{K-1}$ 之后就可以求出期望队长，而华兴对 $G/G/1$ 系统的求解结果不能做到这一点。

1 基本模型

考虑以下排队系统：(1) 顾客到达系统符合参数为 λ 的泊松分布；(2) 系统每隔一段时间 V 产生一次服务，V 的概率密度函数为 $b(t)$；(3) 服务时间足够短，可忽略不计，每次服务都是批量服务，最大批量为 K。

显然，该系统服务的产生不依赖于系统状态，即使系统内无顾客等候，服务也会发生。服务机构就象按照时间表(Time-table) 运行一样，完全独立于系统状态。不妨将此系统称为 $M/G^K/1$ 系统。

2 平稳分布的解析解

设 $P\{V \leq t\} = B(t) = \int_0^t b(t)\mathrm{d}t$，令 N_n 为第 n 批顾客离开之后仍留在系统内的顾客人数，不难证明$\{N_n\}$ 组成一个马尔科夫链。A_n 表示在第 n 批顾客离去时刻与第 $n+1$ 次服务发生时刻之间系统新增的顾客人数。V_n 表示这段时间的长度。这里的$\{V_n\}$ 和$\{A_n\}$ 都是独立同分布的随机变量序列。

$\{A_n\}$ 的概率分布满足：

$$\begin{aligned} \alpha_k = P\{A_n = k\} &= \int_0^{\infty} P\{An = k | V_n = t\} \mathrm{d}B(t) \\ &= \int_0^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \mathrm{d}B(t) \quad (0 \leq k < \infty) \end{aligned}$$

为使平稳分布存在，要求 $\rho = \frac{E[A_n]}{K} = \frac{\sum_{k=0}^{\infty} k\alpha_k}{K} < 1$。我们有以下关系式：

$$N_{n+1} = \begin{cases} N_n + A_n - K & (N_n + A_n > K) \\ 0 & (N_n + A_n \leq K) \end{cases}$$

设 $P_{ij} = P\{N_{n+1} = j | N_n = i\}$，由马尔科夫链的特性，$P_{i,0} = P\{N_{n+1} = 0 | N_n = i\} = P\{A_n \leq K - i\} = \sum_{j=0}^{K-i} \alpha_j (0 \leq i \leq K)$，$P_{j0} = 0\ (j > K)$，$P_{ij} = P\{N_n + 1 = j | N_n = i\} = P\{A_n =$

$K+i-j\} = a_{K+i-j}(j>0, 0 \le i \le K+j)$, $P\{N_n+1=j|N_n=i\}=0(j\ge 0, i>K+j)$,由此得到一步转移概率矩阵:

$$P=\begin{pmatrix} P\{A\le K\} & P\{A=K+1\} & P\{A=K+2\} & \cdots \\ P\{A\le K-1\} & P\{A=K\} & P\{A=K+1\} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ P\{A\le 0\} & P\{A=1\} & P\{A=2\} & \cdots \\ 0 & P\{A=0\} & P\{A=1\} & \cdots \\ 0 & 0 & P\{A=0\} & \cdots \\ 0 & 0 & 0 & \cdots \\ M & M & M & \cdots \end{pmatrix}$$

将 N_n 的稳态概率分布记为 π, $\pi=(\pi_0,\pi_1,\pi_2,\cdots)$,由 $\pi P=\pi$ 得到稳态方程组:

$$\begin{cases} \pi_0 = \sum_{j=0}^{K}\sum_{i=0}^{K-j}\alpha_i\pi_j \\ \pi_i = \sum_{j=0}^{K+i}\alpha_j\pi_{K+i-j} \quad (i>0) \end{cases} \tag{$*$}$$

设:$\prod(s)=\sum_{i=0}^{\infty}\pi_i s^i$, $A(s)=\sum_{i=0}^{\infty}\alpha_i s^i$, $\prod A_k(s)=\sum_{i=0}^{K}\sum_{j=0}^{i}\pi_j\alpha_{i-j}s^i=\pi_0\alpha_0+(\pi_1\alpha_0+\pi_0\alpha_1)s+\cdots+(\pi_K\alpha_0+\pi_{K-1}\alpha_1+\cdots+\pi_0\alpha_K)s^K$ 为$\prod(s)\cdot A(s)$ 的前$(K+1)$ 项。

由稳态方程组($*$)可以得到:

$$\prod(s)=\frac{\prod A_k(s)-\pi_0 s^K}{A(s)-s^K}$$

由于概率母函数$\prod(s)$ 中含有未知量 $\pi_0,\pi_1,\cdots,\pi_K$,下面将它们求出。首先证明 $A(s)-s^K$ 在 $|s|<1$ 内有 $K-1$ 个根:$A(s)$ 是母函数,故在 $|s|\le 1$ 上解析,$A(s)-s^K=(s-1)[-P(s)+Q(s)]$,其中,$P(s)=\sum_{n=0}^{K}\sum_{i=0}^{K-n-1}a_n s^i$, $Q(s)=\sum_{n=K}^{\infty}\sum_{i=0}^{n-K-1}\alpha_n s^i$,在 $|s|=1$ 上,$|P(1)|=\sum_{n=0}^{K}\alpha_n(K-n)$, $|Q(1)|=\sum_{n=K}^{\infty}\alpha_n(n-K)$。而由 $\rho<1\Rightarrow K>E(An)=\sum_{n=0}^{\infty}\alpha_n\cdot n$ 知:

$$K\sum_{n=0}^{\infty}\alpha_n>\sum_{n=0}^{\infty}\alpha_n\cdot n\Rightarrow\sum_{n=0}^{K}\alpha_n(K-n)>\sum_{n=K}^{\infty}\alpha_n(n-K)$$

这样,在 $|s|=1$ 上,$|P(s)|>|Q(s)|$。由 Rouché 定理,$A(s)-s^K=(s-1)(-P(s)+Q(s))$ 在 $|s|<1$ 内有 $K-1$ 个限,加上一个 $s_0=1$,在 $|s|\le 1$ 上有 K 个根。但$\prod(s)$ 是母函数,在 $|s|\le 1$ 内处处解析,因而分子必含有因式项$(s-1)(s-s_1)\cdots(s-s_{K-1})$,否则分母在 $s=s_i(i=0,1,\cdots,K-1)$ 处取 0,$\prod(s)$ 在 $|s|\le 1$ 上将存在极点。由于分子恰好是 K 次多项式,所以恰可写成:$C(s-1)(s-s_1)\cdots(s-s_{K-1})$(其中 C 是一个待定常数)。设 $F(s)=(s-1)(s-s_1)\cdots(s-s_{K-1})$, $G(s)=A(s)-s^K$,母函数可以写成:

$$\prod(s)=\frac{C(s-1)(s-s_1)\cdots(s-s_{K-1})}{A(s)-s^K}=\frac{C\cdot F(s)}{G(s)}$$

由于$\prod(1)=1$,由洛毕达法则:

$$\left.\frac{C\cdot F'(s)}{G'(s)}\right|_{s=1}=1\Rightarrow C=\left.\frac{G'(s)}{F'(s)}\right|_{s=1}$$

因而$\prod(s)$ 可以通过求解 $A(s)-s^K$ 所有单位圆内的根以及 C 加以确定。另外我们还有：

$$A(s)=\sum_{k=0}^{\infty}\int_0^{\infty}\frac{(\lambda t)^k}{k!}e^{-\lambda t}s^k\mathrm{d}B(t)=\int_0^{\infty}e^{-(\lambda-\lambda s)t}\mathrm{d}B(t)=B^*(\lambda-\lambda s)$$

例 1 服务间隔服从负指数分布，即 $b(t)=\mu e^{\mu t}$，满载量 1 人，即 $K=1$。

解 由 $b(t)=\mu e^{\mu t}$ 可以推出：$B^*(s)=\dfrac{\mu}{\mu+s}$，则：$A(s)=B^*(\lambda-\lambda s)=\dfrac{\mu}{\mu+\lambda-\lambda s}$。

$$\prod(s)=\frac{\pi_0\alpha_0+(\pi_0\alpha_1+\pi_1\alpha_0)s-s}{\dfrac{\mu}{\lambda+\mu-\lambda s}-s}$$

显然分母在单位圆上只有一个根 $s=1$，分子可以写成 $C(s-1)$，由洛毕达法则可得：$C=\rho-1$。其中 $\rho=\lambda/\mu$，这样由：

$$\prod(s)=\frac{(\rho-1)(s-1)}{\mu/(\lambda+\mu-\lambda s)-s}=(\rho-1)\sum_{k=1}^{\infty}(\rho s)^k-(\rho-1)(\rho+1)\sum_{k=0}^{\infty}(\rho s)^k$$

$$=(1-\rho^2)-(\rho-1)\sum_{k=1}^{\infty}\rho^{k+1}s^k$$

可以得到：$\pi_0=1-\rho^2,\pi_k=(1-\rho)\rho^{k+1},k\geq 1$

例 2 若服务间隔服从单点分布，即：服务间隔定长，$b(1/\mu)=1,b(t)=0(t\neq 1/\mu)$，满载 1 人，即 $K=1$。

解 容易算出：$\alpha_k=\dfrac{(\lambda/\mu)^k}{k!}e^{-\lambda/\mu}$，$A(s)=e^{-\rho(1-s)}$，分子同样可以写成 $C(s-1)$，由洛毕达法则：$C=\rho-1$，这样 $\pi_0=e^{\rho}(1-\rho),\pi_1=e^{\rho}(e^{\rho}-\rho-1)(1-\rho)$。

例 3 服务间隔服从负指数分布 $b(t)=\mu e^{-\mu t}$，满载 K 人。

解 $$\prod(s)=\frac{\prod_K A(s)-\pi_0 s^K}{A(s)-s^K}=\frac{(\lambda+\mu-\lambda s)(\prod_K A(s)-\pi_0 s^K)}{\mu-s^K(\lambda+\mu-\lambda s)}$$

因为分母恰好有 K 个在单位圆内的根，因此分子也应该有与之相同的 K 个根。显然分子有一个根$(\lambda+\mu)/\lambda>1$，而$\prod_K A(s)-\pi_0 s^K$ 是 K 次多项式，恰有 K 个根，因此可以与分母相互消掉一个 K 次多项式，剩下一个常数。分母是 $K+1$ 次多项式，剩下$(s-s_0)$ 项，其中 s_0 是分母 $K+1$ 次多项式在单位圆之外的一个根。故利用洛毕达法则：

$$\text{原式}=\frac{C(\lambda+\mu-\lambda s)}{s-s_0}=\frac{(1-s_0)(\rho+1-\rho s)}{s-s_0}=\frac{(s_0-1)(\rho+1-\rho s)}{(1-s/s_0)s_0}$$

$$=\frac{s_0-1}{s_0}(\rho+1-\rho s)\sum_{k=0}^{\infty}\left(\frac{s}{s_0}\right)^k=\frac{s_0-1}{s_0}(\rho+1)\sum_{k=0}^{\infty}\left(\frac{s}{s_0}\right)^k-\frac{s_0-1}{s_0}\cdot\rho\sum_{k=1}^{\infty}\frac{s^k}{s_0^{k-1}}$$

$$\Rightarrow\pi_0=\frac{s_0-1}{s_0}(\rho+1),\pi_k=\frac{s_0-1}{s_0}(\rho+1)\frac{1}{s_0^k}-\frac{s_0-1}{s_0}\rho\frac{1}{s_0^{k-1}}=\frac{s_0-1}{s_0}\left(\frac{\rho+1}{s_0}-\rho\right)$$

3 平均等待时间和平均队列长度：

如前所述： $$N_{n+1}=\max(0,N_n+A_n-K)$$

记 $$\delta_n=\begin{cases}0 & (if\ N_n+A_n-K\geq 0)\\ -(N_n+A_n-K) & (if\ N_n+A_n-K\leq 0)\end{cases}$$

则 $$N_{n+1}=(N_n+A_n-K)+\delta_n\Rightarrow(N_{n+1}-\delta_n)^2=(N_n+A_n-K)^2,$$

记 $U_n=A_n-K$，展开并注意到 $N_{n+1}\delta_n=0$，我们有 $N_{n+1}^2+\delta_n^2=N_n^2+U_n^2+2N_nU_n$，注意到 N_n

与 A_n 无关，从而与 U_n 无关，两边取期望并让 $n \to \infty$（因而 $E(N_{n+1}^2) = E(N_n^2)$ 得：$E(N) = \dfrac{E(\delta^2) - E(U^2)}{2E(U)}$，其中 $E(U^2) = Var(U) + (E(U))^2 = Var(A) + (K - E[A])^2$，$\delta_n$ 可解释为每趟列车经过时闲置的座位，因而 $E[\delta^2] = E[\delta^2 | N_n + A_n - K \leq 0] = E[(N_n + A_n - K)^2 | N_n + A_n - K \leq 0] = K^2 \cdot \pi_0\alpha_0 + (K-1)^2(\pi_0\alpha_1 + \pi_1\alpha_0) + \cdots + (\pi_0\alpha_{K-1} + \pi_1\alpha_{K-2} + \cdots + \pi_{K-1}\alpha_0)$。可见，只要求出 $\pi_0, \pi_1, \cdots, \pi_{K-1}$ 就可以得到 $E[\delta^2]$ 之值，进而可求 $E[N]$。但是要求 $\pi_0, \pi_1, \cdots, \pi_{K-1}$，需要求解 $A(s) - s^K$ 单位圆内的根，一般情况下，找到这些解的解析形式是十分困难的。如果精度要求不高，我们可以通过计算机模拟找到 $\pi_0, \pi_1, \cdots, \pi_{K-1}$ 的近似值，代入式中可以求出 $E(N)$。

例 4　条件同例 1，求平均队长。

解

$$A(s) = \frac{\mu}{\lambda + \mu - \lambda s} \Rightarrow A'(s)|_{s=1} = E(A) = \rho, A''(s)|_{s=1} + A'(s)|_{s=1} = E(A^2) = 2\rho^2 + \rho$$

$$Var(A) = E(A^2) - [E(A)]^2 = \rho^2 + \rho, E(A) = \rho,$$

$$E[\delta^2] = \pi_0\alpha_0 = (1-\rho^2) \cdot \frac{\mu}{\lambda + \mu} = \frac{(1-\rho^2)}{(1+\rho)} = 1 - \rho,$$

$$E(N) = \frac{E(\delta^2) - E(U^2)}{2E(U)} = \frac{1-\rho-\rho^2-\rho-1-\rho^2+2\rho}{2(\rho-1)} = \frac{-2\rho^2}{2(\rho-1)} = \frac{\rho^2}{1-\rho}。$$

例 5　条件同例 2，求平均队长。

解　$Var(A) = \rho, E(A) = \rho, E[\delta^2] = e^{\rho}(1-\rho) \cdot e^{-\rho} = 1 - \rho,$

$$E(N) = \frac{E(\delta^2) - E(U^2)}{2E(U)} = \frac{1-\rho-\rho-1-\rho^2+2\rho}{2(\rho-1)} = \frac{-\rho^2}{2(\rho-1)} = \frac{\rho^2}{2(1-\rho)}。$$

4　结束语

当服务间隔满足负指数分布时，批量服务系统的解析解可以求出。但如果不满足负指数分布，解析解难以求出，困难在于 $A(s) - s^K = 0$ 在单位圆内的所有解的求解。但本文给出了求解这类服务系统的方向，在一些特殊情况下，$A(s) - s^K = 0$ 在单位圆内的解还是可以通过数值迭代的方法得到近似解。也可以通过连乘转移概率矩阵的方法得到稳态概率分布的近似解。

对于这类复杂服务系统的研究，计算机仿真也是一种有效的方法。但计算机仿真容易产生累计误差，在 λ 接近 1 时，误差就更为明显。不过计算机仿真结果也可以作为对数值迭代方法求解的结果的一个参考。摄动法也可用于这类系统的求解，它可以给出解的取值范围，将这种思想用于计算机仿真，可以检测仿真结果的波动范围和可靠性。

参考文献

[1] Leonard Kleinrock. Queueing systems[M]. New York: John Wiley & Sons Inc., 1975.
[2] 徐光辉. 随机服务系统[M]. 北京：科学出版社，1980.
[3] J. Medhi. Stochastic models in queueing theory[M]. New York: A cademic Press Inc., 1994.
[4] J. W. Cohen. The single server queue[M]. New York: North-Holland Publish Company, 1982.
[5] 华兴. 排队论与随机服务系统[M]. 上海：上海翻译出版社，1987.