

基于可变速率 $M/M/c$ 排队的接触匹配模型分析*

谭玉廷, 徐秀丽

(燕山大学 理学院, 河北 秦皇岛 066004)

摘要:【目的】针对实际生活中利用手机软件呼叫出租车的排队场景,建立并分析了可变速率及可变服务率的接触匹配的 $M/M/c$ 排队模型。【方法】利用概率分析方法导出拟生灭过程的状态转移规律及无穷小生成元矩阵,利用矩阵几何解方法给出系统的稳态平衡条件、稳态概率分布及系统的主要性能指标。【结果】通过数值分析讨论系统参数对于性能指标的影响,并建立系统收益函数研究系统最优收益。【结论】得到了系统参数对系统性能指标及系统收益的影响关系,表明在接触匹配中应当同时调整服务率及接触匹配成功率才能显著提高系统收益。

关键词: 排队模型;可变速率;可变服务率;接触匹配

中图分类号: O226

文献标志码: A

文章编号: 1672-6693(2022)06-0028-11

随着信息时代的迅速发展,生活中的排队场景也发生了许多变化,这也促使排队的模型理论进一步发展。生活中的实际排队模型,越来越将顾客以及服务台的个人感受考虑到模型中。在 20 世纪 50 年代,Haigh^[1]提出了带有阻塞的 $M/M/1$ 排队模型,从顾客行为角度来分析排队模型。类似地,也可从服务台角度来研究排队模型,这往往与系统收益有关,田乃硕^[2]提出了由系统中顾客数控制休假终止时间的 $M/M/1$ 服务员休假排队模型,讨论了以净收益为最大目标的休假策略优化问题。

Bai 等人^[3]提出了一个内生供给(参与代理的数量)和内生需求(客户请求率)的分析模型来研究按需服务平台,在客户效用函数中加入了基于排队模型的稳态等待时间性能,以描述平台利润最大化的最优价格和工资率。Sudtachat 等人^[4]提出了易腐的库存模型,将易腐的库存模型系统发展成一个带有不耐烦顾客的马尔可夫过程。Ashour 等人^[5]提出了一个分析框架,用以估计服务率变化的优先级排队系统的队列长度等系统性能指标。卢宜玲等人^[6]提出了疫情下医生因为疲劳度等问题会使得医院的服务率降低,进而提出了通过可变服务率来使得医生的疲劳度最小化及患者的满意度最大化。Wang 等人^[7]考虑在多服务器忙期排队系统中可能会出现阻塞和拒绝的顾客的不耐烦行为,并研究了服务提供商在系统繁忙时试图提高服务率的策略行为,即等待队列到达一个阈值时,顾客到达率会相应减少,系统为了系统收益会提高服务率使得等待队列小于阈值,从而减少顾客损失。Ke 等人^[8]提出了带休假的 $M/M/c$ 阻塞重试队列,即顾客进入系统没有空闲服务台时顾客可以选择离开系统也可以继续等待,离开系统再次进入重试轨道的顾客的速率会根据等待队列的大小而变化。孙敬煊等人^[9]则提出了接触匹配的排队理论,即顾客与服务台之间在接受服务之前,会存在一个接触匹配的过程,即顾客可以选择离开系统,服务台也可以根据实际情况拒绝与顾客进行匹配。

近些年来,随着网络性能的发展以及智能手机的普及,网络打车逐渐成为人们的日常出行方式之一。在网络打车这个概念之前,生活中与之类似的是网络租车,George 等人^[10]讨论了汽车租赁公司控制公司最佳车队规模的问题,提出了一个基于稳态的优化问题来确定利润最优的车队规模,并给出了一个求解精确解的迭代算法。在实际生活中出租车在对顾客进行服务前,会存在一定的匹配时间,Shi 等人^[11]研究了非零匹配时间的客运出租车排队系统,考虑了乘客与出租车司机逗留时间的尾部分布,来对相关系统性能指标进行数值模拟。由于空间异质性等因素,网络打车系统会在某一时间段的某一区域出现大量顾客,此时系统需要一定的激励措施来增加系统收益,Qian 等人^[12]研究了出租车运营的激励策略,在最优分配下实现效率最优化。Zhang 等人^[13]提出了竞争性的出租车市场的创新建模结构,将出租车系统内使用类似 $M/M/1$ 排队系统来匹配顾客出行,城市道路系统

* 收稿日期:2022-05-19 修回日期:2022-08-19 网络出版时间:2022-12-08 14:31

资助项目:国家自然科学基金(No. 62171143);河北省高等学校科学研究重点项目(No. ZD2019079)

第一作者简介:谭玉廷,男,研究方向为排队论与随机过程,Email:1176878197@qq.com;通信作者:徐秀丽,女,教授,博士生导师,Email:xxl-ysu@163.com

网络出版地址:https://kns.cnki.net/kcms/detail//50.1165.N.20221207.1716.013.html

使用多服务台的 M/M/c 排队队列进行建模,将道路网络状态反馈给出租车系统来进行出租车最优分配及系统效益最优化。

本文在上述研究的基础上,考虑这样一个问题:在一个固定的区域内,按需服务平台一方面为代理商(出租车司机)提供顾客信息进而进行服务,另一方面为有出行需求的顾客提供出租车信息从而提高顾客的打车效率。与传统排队模型不同的是,服务台和顾客是同时加入系统,且顾客与服务台在进行服务前有一个接触匹配的过程,即接触匹配成功再进行服务,不成功则顾客离开系统。本文将顾客和服务台都考虑在内,在 M/M/c 排队模型的基础上,加入了顾客中途离开、接触匹配、可变到达率和可变服务率等策略,并给出了相关的稳态分布求解方法和相关系统性能指标,最后通过数值模拟讨论了顾客到达率与服务率对于系统性能指标以及系统收益函数的影响。

1 模型假设

本文在 M/M/c 排队模型的基础上加入了可变到达率、可变服务率以及接触匹配的服务机制。顾客进入系统后若有空闲的服务台,顾客与空闲的服务台直接进行接触匹配,若接触匹配成功则顾客接受服务,服务台处于忙期状态;若接触匹配不成功,顾客则可以离开系统,也可以在等待队列中继续等待与新的服务台进行接触匹配。若顾客进入系统后没有空余的服务台,顾客可以选择在等待队列中等待,也可以选择离开系统,顾客在等待队列中等待新的服务台加入系统进行接触匹配或者等待系统中的服务台处于闲期时接受服务。当等待队列 $n \geq c_1$ 时,顾客考虑到进入系统后等待时间会过长,则进入系统的速率会随着等待队列的变大而减小。与此相同的是,服务台考虑到系统收益,服务率会随着等待队列的变大而增大。到达率和服务率均可变的接触匹配的 M/M/c 排队模型的具体描述如下:

1) 当顾客的等待队列 $0 \leq n < c_1$ 时,顾客的到达间隔时间与服务时间分别服从参数为 λ, μ 的指数分布,即顾客的到达率为 λ ,单个服务台的服务率为 μ ,服务规则为先到先服务(First-come first-served)。

2) 当顾客的等待队列 $c_1 \leq n \leq N$ 时,顾客的到达率会发生一定的变化,由于等待队列较长,顾客进入系统的速率会随着等待队列的长度增大而不断衰减。此时顾客动态到达率为 $\lambda_n = \lambda \alpha_n$,其中:

$$\alpha_n = \frac{1}{2} \left(\ln \frac{N}{c_1} \right)^{-1}, \alpha_n = \begin{cases} 1 & 2 \leq n < c_1 \\ 1 + a_1 \ln \frac{c_1}{n} & c_1 \leq n \leq N \end{cases}, \text{即此时到达率满足 } \lambda_{\min} = 0.5\lambda。$$

3) 当顾客的等待队列 $c_1 \leq n \leq N$ 时顾客的到达率会不断减小,同时正处于工作状态的服务台为了增加系统收益选择加大服务速率,则每个工作的服务台的动态服务率为 $\mu_n = \mu \beta_n$,其中:

$$\beta_n = \frac{1}{2} \left(\ln \frac{c_1}{N} \right)^{-1}, \beta_n = \begin{cases} 1 & 2 \leq n < c_1 \\ 1 + b_n \ln \frac{c_1}{n} & c_1 \leq n \leq N \end{cases}, \text{即此时服务率满足 } \mu_{\max} = 1.5\mu。$$

4) 顾客的主动匹配和服务台的拒绝匹配,视为两个独立的过程,且均服从指数分布。顾客的主动匹配成功率为 η ,服务台的拒绝匹配率为 ϵ 。处于等待队列的顾客,会出现等待不耐烦情况导致中途离开等待队列,顾客会以 σ 的速率离开系统,且服从指数分布,当等待队列 $n \geq N$ 时,顾客离开系统率达到阈值,此时 $\sigma_n = N\sigma$;与此类似的是,当系统出现等待队列时,所有的服务台都处于忙期或者接触匹配期,服务台会以速率 q 进入系统直接与队列中第一个等待顾客进行接触匹配,服务台进入系统与顾客中途离开队列是相互独立的过程。系统内最多有 c 个服务台处于忙期或者接触匹配的状态。特别的是,当系统中没有服务台处于闲期时,顾客会以概率 b 进入等待队列,顾客会以概率 $1-b$ 不选择等待而直接离开系统,即此时顾客的到达率为 $b\lambda_n$ 。

5) 当等待队列 $n > N$ 时,顾客的到达率和服务率均保持不变,此时 $\lambda_n = 0.5\lambda, \mu_n = 1.5\mu$ 。

2 模型建立

为了进行建模分析,设 $K(t)$ 为 t 时刻系统中处于忙期的服务台的个数, $W(t)$ 为 t 时刻等待队列的人数, $S(t)$ 为 t 时刻系统中处于接触匹配的服务台的个数,则三维马尔科夫过程 $\{W(t), K(t), S(t), t \geq 0\}$ 是一个拟生灭过程,它的状态空间为 $\Omega = \{(n, k, s) | 0 \leq k + s \leq c, n \geq 0\}$ 。将状态按字典序排列,它的无穷小生成元矩阵可以写成下列分块三对角形式:

$$Q = \begin{bmatrix} A_0 & C_0 & & & & \\ B_1 & A_1 & C_0 & & & \\ & \ddots & \ddots & \ddots & & \\ & & B_{c_1-1} & A_{c_1-1} & C_0 & \\ & & & B_{c_1} & A_{c_1} & C_{c_1} \\ & & & & \ddots & \ddots \\ & & & & & B_{N-1} & A_{N-1} & C_{N-1} \\ & & & & & & B_N & A_N & C_N \\ & & & & & & & B_N & A_N & C_N \\ & & & & & & & & B_N & A_N & C_N \\ & & & & & & & & & \ddots & \ddots & \ddots \end{bmatrix},$$

其中各子块矩阵如下:

1) A_0 为 $\frac{(c+2) \times (c+1)}{2}$ 阶方阵, 可表示为:

$$A_0 = \begin{bmatrix} Y_0^0 & X_0^0 & & & & \\ Z_1^0 & Y_1^0 & X_1^0 & & & \\ & Z_2^0 & Y_2^0 & X_2^0 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ & & & & Z_{c-1}^0 & Y_{c-1}^0 & X_{c-1}^0 \\ & & & & & Z_c^0 & Y_c^0 \end{bmatrix},$$

其中: Y_k^0 为 $k+1$ 阶方阵, 可记为:

$$Y_0^0 = -(\lambda + b\lambda), Y_k^0 = \begin{bmatrix} h_0 & l_0 & & & \\ & h_1 & l_1 & & \\ & & \ddots & \ddots & \\ & & & h_{k-1} & l_{k-1} \\ & & & & h_k \end{bmatrix}, 1 \leq k \leq c,$$

且当 $1 \leq k \leq c-1$ 时,

$$h_i = -[\lambda + b\lambda + i\mu + (k-i)(\epsilon + \eta)], 0 \leq i \leq k, \\ l_i = (k-i)\eta, 0 \leq i \leq k-1,$$

当 $k=c$ 时,

$$h_i = -[b\lambda + i\mu + (k-i)(\epsilon + \eta)], 0 \leq i \leq k, \\ l_i = (k-i)\eta, 0 \leq i \leq k-1.$$

X_k^0 为 $(k+1) \times (k+2)$ 阶矩阵, 可表示为:

$$X_0^0 = [\lambda, 0], X_k^0 = \begin{bmatrix} \lambda & & & 0 \\ & \lambda & & 0 \\ & & \ddots & \vdots \\ & & & \lambda & 0 \end{bmatrix}, 1 \leq k \leq c-1.$$

Z_k^0 为 $(k+1) \times k$ 阶矩阵, 可记为:

$$Z_1^0 = \begin{bmatrix} \epsilon \\ \mu \end{bmatrix}, Z_k^0 = \begin{bmatrix} p_0 & & & \\ r_1 & p_1 & & \\ & \ddots & \ddots & \\ & & r_{k-1} & p_{k-1} \\ & & & r_k \end{bmatrix}, 2 \leq k \leq c,$$

且当 $2 \leq k \leq c$ 时,

$$\begin{aligned} p_i &= (k-i)\varepsilon, 0 \leq i \leq k-1, \\ r_i &= i\mu, 1 \leq i \leq k. \end{aligned}$$

2) 矩阵 $\mathbf{C}_0$ 为 $\frac{(c+2) \times (c+1)}{2}$ 阶方阵, 可表示为:

$$\mathbf{C}_0 = \begin{bmatrix} c_{0,0} & & & \\ & \mathbf{c}_{0,1} & & \\ & & \ddots & \\ & & & \mathbf{c}_{0,c} \end{bmatrix},$$

其中 $i+1$ 阶方阵 $\mathbf{c}_{0,i} = \begin{bmatrix} b\lambda & & \\ & b\lambda & \\ & & \ddots \\ & & & b\lambda \end{bmatrix}, 0 \leq i \leq c$ 。

3) $\mathbf{A}_n (1 \leq n \leq c_1)$ 可表示为:

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{Y}_0^n & \mathbf{X}_0^n & & & \\ \mathbf{Z}_1^n & \mathbf{Y}_1^n & \mathbf{X}_1^n & & \\ & \mathbf{Z}_2^n & \mathbf{Y}_2^n & \mathbf{X}_2^n & \\ & & \ddots & \ddots & \ddots \\ & & & \mathbf{Z}_{c-1}^n & \mathbf{Y}_{c-1}^n & \mathbf{X}_{c-1}^n \\ & & & & \mathbf{Z}_c^n & \mathbf{Y}_c^n \end{bmatrix},$$

其中 $\mathbf{Y}_k^n$ 为 $k+1$ 阶方阵, 可记为:

$$\mathbf{Y}_0^n = -(\lambda + b\lambda + n\sigma + q), \mathbf{Y}_k^n = \begin{bmatrix} h_{n,0} & l_0 & & & \\ & h_{n,1} & l_1 & & \\ & & \ddots & \ddots & \\ & & & h_{n,k-1} & l_{k-1} \\ & & & & h_{n,k} \end{bmatrix}, 1 \leq k \leq c,$$

且当 $1 \leq k \leq c-1$ 时,

$$\begin{aligned} h_{n,i} &= -[\lambda + b\lambda + i\mu + (k-i)(\varepsilon + \eta) + n\sigma + q], 0 \leq i \leq k, \\ l_i &= (k-i)\eta, 0 \leq i \leq k-1, \end{aligned}$$

当 $k=c$ 时,

$$\begin{aligned} h_{n,i} &= -[b\lambda + i\mu + (k-i)(\varepsilon + \eta) + n\sigma], 0 \leq i \leq k, \\ l_i &= (k-i)\eta, 0 \leq i \leq k-1. \end{aligned}$$

同理, $\mathbf{X}_k^n$ 为 $(k+1) \times (k+2)$ 阶矩阵, 可记为:

$$\mathbf{X}_0^n = [\lambda, 0], \mathbf{X}_k^n = \begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{bmatrix}, 1 \leq k \leq c-1.$$

$\mathbf{Z}_k^n$ 为 $(k+1) \times k$ 阶矩阵, 可记为:

$$\mathbf{Z}_1^n = \begin{bmatrix} \varepsilon \\ \mu \end{bmatrix}, \mathbf{Z}_k^n = \begin{bmatrix} p_0 & & & \\ r_1 & p_1 & & \\ & \ddots & \ddots & \\ & & r_{k-1} & p_{k-1} \\ & & & r_k \end{bmatrix}, 2 \leq k \leq c,$$

且当 $2 \leq k \leq c$ 时,

$$\begin{aligned} p_i &= (k-i)\epsilon, 0 \leq i \leq k-1, \\ r_i &= i\mu, 1 \leq i \leq k. \end{aligned}$$

4) 矩阵 $B_n (1 \leq n \leq N)$ 可表示为:

$$B_n = \begin{bmatrix} v_{0,n} & d_0 & & & \\ & v_{1,n} & d_1 & & \\ & & \ddots & \ddots & \\ & & & v_{c-1,n} & d_{c-1} \\ & & & & v_{c,n} \end{bmatrix},$$

其中: $v_{k,n}$ 为 $k+1$ 阶矩阵, 可记为:

$$v_{0,n} = n\sigma, v_{k,n} = \begin{bmatrix} n\sigma & & & \\ & n\sigma & & \\ & & \ddots & \\ & & & n\sigma \end{bmatrix}, 1 \leq k \leq c,$$

d_k 为 $k \times (k+1)$ 阶矩阵, 可记为:

$$d_0 = [q \ 0], d_k = \begin{bmatrix} q & & & 0 \\ & q & & 0 \\ & & \ddots & \vdots \\ & & & q & 0 \end{bmatrix}, 1 \leq k \leq c-1.$$

5) $A_n (c_1+1 \leq n \leq N)$ 可表示为:

$$A_n = \begin{bmatrix} Y_0^n & X_0^n & & & \\ Z_1^n & Y_1^n & X_1^n & & \\ & Z_2^n & Y_2^n & X_2^n & \\ & & \ddots & \ddots & \ddots \\ & & & Z_{c-1}^n & Y_{c-1}^n & X_{c-1}^n \\ & & & & Z_c^n & Y_c^n \end{bmatrix},$$

其中: Y_k^n 为 $k+1$ 阶方阵, 可记为:

$$Y_0^n = -(\lambda_n + b\lambda_n + n\sigma + q), Y_k^n = \begin{bmatrix} h_{n,0} & l_0 & & & \\ & h_{n,1} & l_1 & & \\ & & \ddots & \ddots & \\ & & & h_{n,k-1} & l_{k-1} \\ & & & & h_{n,k} \end{bmatrix}, 1 \leq k \leq c,$$

且当 $1 \leq k \leq c-1$ 时,

$$\begin{aligned} h_{n,i} &= -[\lambda_n + b\lambda_n + i\mu_n + (k-i)(\epsilon + \eta) + n\sigma + q], 0 \leq i \leq k, \\ l_i &= (k-i)\eta, 0 \leq i \leq k-1, \end{aligned}$$

当 $k=c$ 时,

$$\begin{aligned} h_{n,i} &= -[b\lambda_n + i\mu_n + (k-i)(\epsilon + \eta) + n\sigma], 0 \leq i \leq k, \\ l_i &= (k-i)\eta, 0 \leq i \leq k-1. \end{aligned}$$

同理, X_k^n 为 $(k+1) \times (k+2)$ 阶矩阵, 可记为:

$$X_0^n = [\lambda_n, 0], X_k^n = \begin{bmatrix} \lambda_n & & & 0 \\ & \lambda_n & & 0 \\ & & \ddots & \vdots \\ & & & \lambda_n & 0 \end{bmatrix}, 1 \leq k \leq c-1.$$

Z_k^n 为 $(k+1) \times k$ 阶矩阵,可记为:

$$Z_1^n = \begin{bmatrix} \epsilon \\ \mu_n \end{bmatrix}, Z_k^n = \begin{bmatrix} p_0 & & & & \\ r_{n,1} & p_1 & & & \\ & \ddots & \ddots & & \\ & & r_{n,k-1} & p_{k-1} & \\ & & & & r_{n,k} \end{bmatrix}, 2 \leq k \leq c,$$

当 $2 \leq k \leq c$ 时,

$$p_i = (k-i)\epsilon, 0 \leq i \leq k-1; r_{n,i} = i\mu_n, 1 \leq i \leq k.$$

6) 矩阵 $C_n (c_1 \leq n \leq N)$ 可表示为:

$$C_n = \begin{bmatrix} c_{n,i} & & & \\ & c_{n,i} & & \\ & & \ddots & \\ & & & c_{n,c} \end{bmatrix},$$

其中 $i+1$ 阶方阵 $c_{n,i} = \begin{bmatrix} b\lambda_n & & & \\ & b\lambda_n & & \\ & & \ddots & \\ & & & b\lambda_n \end{bmatrix}, 0 \leq i \leq c.$

3 稳态分析

在确定系统的稳态性能指标前,要进行系统稳态性的验证,以确定系统稳态分布存在。令 $F = C_N + A_N + B_N$, $e = [e_0, e_1, \dots, e_i, \dots, e_c]^T$, 其中 e_i 为 $i+1$ 维全 1 列向量。记 P 为 F 的平稳概率向量,且 $P = [P_0, P_1, \dots, P_i, \dots, P_c]$, 其中 $P_i = [p_{0,i}, p_{1,i-1}, \dots, p_{i-1,1}, p_{i,0}]$ 。

由 $PF = 0, Pe = 1$ 可得到以下方程组:

当 $i=0$ 时,

$$(\lambda_N + b\lambda_N - N\sigma)p_{0,0} = \epsilon p_{0,1} + \mu_N p_{1,0},$$

当 $1 \leq i \leq c-1$ 时,

$$(\lambda_N + b\lambda_N - N\sigma)(p_{0,i} + p_{1,i-1} + \dots + p_{i,0}) + (b\lambda_N - N\sigma - q)(p_0 + p_1 + \dots + p_{i-1}) = (i+1)\epsilon p_{0,i+1} + \mu_N p_{1,i} + i\epsilon p_{1,i} + 2\mu_N p_{2,i} + \dots + (i+1)\mu_N p_{i+1},$$

当 $i=c$ 时,

$$(\lambda_N + q)(p_{0,c-1} + p_{1,c-2} + \dots + p_{c-1,0}) - (\lambda_N + b\lambda_N - N\sigma)(p_{0,c} + p_{1,c-1} + \dots + p_{c,0}) = c\epsilon p_{0,c} + \mu_N p_{1,c-1} + (c-1)\epsilon p_{1,c-1} + 2\mu_N p_{2,c-2} + \dots + c\mu_N p_{c,0}.$$

此外,

$$\sum_{i=0}^c P_i e_i = 1.$$

求解上述方程组可知 P 满足 $PB_N e > PC_N e$, 由文献[14]的定理 2.5 可知系统的稳态分布存在。记 $\pi_{k,s}^n$ 表示系统稳态概率,即稳态下系统中有 k 个处于忙期的服务台, s 个处于接触匹配的服务台, n 个处于等待队列的顾客的概率, $0 \leq k+s \leq c, n \geq 0$ 。记稳态概率向量 $\pi = [\pi_0, \pi_1, \dots, \pi_c, \dots]$, 其中 $\pi_n = [\pi_{0,0}^n, \pi_{0,1}^n, \pi_{1,0}^n, \pi_{0,2}^n, \pi_{1,1}^n, \pi_{2,0}^n, \dots, \pi_{0,c}^n, \pi_{1,c-1}^n, \dots, \pi_{c,0}^n]$, 从而可以得到下列平衡方程组:

$$\pi_0 A_0 + \pi_1 B_1 = 0, \quad (1)$$

$$\pi_{n-1} C_0 + \pi_n A_n + \pi_{n+1} B_n = 0, 1 \leq n \leq c_1, \quad (2)$$

$$\pi_{n-1} C_{n-1} + \pi_n A_n + \pi_{n+1} B_n = 0, c_1+1 \leq n \leq N-1, \quad (3)$$

$$\pi_{N-1} C_{N-1} + \pi_N A_N + \pi_N B_N = 0, \quad (4)$$

$$\pi_N R^{n-1-N} C_N + \pi_N R^{n-N} A_N + \pi_N R^{n-N+1} B_N = 0, n \geq N+1, \quad (5)$$

$$\sum_{n=0}^{\infty} \pi_n e = 1. \quad (6)$$

直接求解上述方程组比较困难,为了求解稳态分布,由拟生灭过程和矩阵几何解方法可知,需求解率阵 R ,即矩阵方程 $R^2 B_N + R A_N + C_N = 0$ 的最小非负解,且使得谱半径 $SP(R) < 1$ 。由于直接解二次矩阵方程得到率阵 R 的精确解是十分困难的,下面采用 Neuts^[15] 给出的数值迭代方法来近似计算,具体迭代过程为:

- 1) 确定系统参数 $c, c_1, N, \lambda, \mu, \varepsilon, \eta, \sigma, q, b$ 以及精度 $\delta = 10^{-10}$ 。
 - 2) 输入 C_N, B_N, A_N 。
 - 3) 定义 $R(n) = R, R = -(C_N + R^2 B) A_N^{-1}, R(n+1) = R$ 。
 - 4) 若 $\|R(n+1) - R(n)\| < \delta$ 成立,则取 $R = R(n+1)$,迭代结束。否则转步骤 3。
- 通过上述算法求解出率阵 R 后,根据(1)~(4)式可以得出下列递推方程:

$$\pi_0 = \pi_1 B_1 (-A_0)^{-1} = \pi_1 \omega_1, \quad (7)$$

$$\pi_n = \pi_{n+1} B_n [-(\omega_n C_0 + A_n)]^{-1} = \pi_{n+1} \omega_{n+1}, 1 \leq n \leq c_1, \quad (8)$$

$$\pi_n = \pi_{n+1} B_n [-(\omega_n C_n + A_n)]^{-1} = \pi_{n+1} \omega_{n+1}, c_1 + 1 \leq n \leq N-1, \quad (9)$$

$$\pi_N \omega_N C_N + \pi_N A_N + \pi_N R B_N = 0, \quad (10)$$

由拟生灭过程和矩阵几何解可知:

$$\pi_{N+n_1} = \pi_N R^{N+n_1}, n_1 > 0. \quad (11)$$

结合(7)~(11)式并根据(6)式得到:

$$1 = \sum_{n=0}^{\infty} \pi_n e = \pi_N \left[\sum_{n=1}^N \prod_{i=n}^N \omega_i + (I - R)^{-1} \right] e. \quad (12)$$

联立(7)~(12)式可以求得稳态概率向量 $\pi = (\pi_0, \pi_1, \pi_2, \dots, \pi_N, \dots)$ 。

利用系统稳态分布可进一步给出系统性能指标,为了表示方便,令 L_q 表示稳态下系统等待队列长度, C_s 表示稳态下处于接触匹配状态的服务台数, B_s 表示稳态下处于忙期的服务台数,则有:

- 1) 稳态下系统平均等待队长:

$$E(L_q) = \sum_{j=1}^{N-1} j \pi_j e + N \pi_N R e + (N+1) \pi_N R^2 e + \dots = \pi_N \left[\sum_{j=2}^N (j-1) \prod_{i=j}^N \omega_i + N(I - R)^{-1} + R(I - R)^{-2} \right] e.$$

- 2) 稳态下处于接触匹配的平均服务台数:

$$E(C_s) = \sum_{j=0}^{\infty} \pi_j t_1 = \pi_N \left[\sum_{n=1}^N \prod_{i=n}^N \omega_i + (I - R)^{-1} \right] t_1,$$

其中, $t_1 = [0, 1, 0, 2, 1, 0, \dots, c, c-1, \dots, 0]^T$ 。

- 3) 稳态下处于忙期的平均服务台数:

$$E(B_s) = \sum_{j=0}^{\infty} \pi_j t_2 = \pi_N \left[\sum_{n=1}^N \prod_{i=n}^N \omega_i + (I - R)^{-1} \right] t_2,$$

其中, $t_2 = [0, 0, 1, 0, 1, 2, \dots, 0, 1, \dots, c]^T$ 。

4 数值分析

在实际生活中,由于节假日以及上下班等社会行为会极大影响顾客到达率 λ ,接触匹配成功率 η 等系统参数的变化。人们由于节假日或上下班高峰期等因素,出行意愿较大,对订单价格的敏感度会下降,导致 η 会增大。当系统中顾客数量较多时,服务台进入系统的意愿也会加大,同时系统为了更好的运营也会采取一定的激励措施来鼓励服务台与顾客进入系统,服务台与顾客的接触匹配率增加。这些行为都会使得系统参数发生变化,系统的性能指标和系统收益也会随着系统参数的变化而变化。下面给出数值例子来刻画性能指标和系统收益函数随系统参数的趋势变化以及实际意义。

4.1 系统性能指标的敏感性分析

为了方便直观地分析系统参数变化对系统性能指标的影响,下面假设 $c=3, c_1=7, N=13, b=0.5, \sigma=$

0.3,研究顾客到达率 λ 、服务率 μ 、接触匹配率 η 、服务台拒绝率 ϵ 和服务台进入系统率 q 对系统性能指标的影响。

当 $q=1, \mu=0.8, \epsilon=0.3, \eta=0.3$ 时,图1反映了 $E(L_q)$ 、 $E(C_s)$ 和 $E(B_s)$ 随 λ 的变化趋势。可以发现随着 λ 的增大, $E(L_q)$ 会逐渐增加,而 $E(C_s)$ 、 $E(B_s)$ 却增长十分缓慢。这是因为 λ 越大,进入系统的顾客越来越多,而 η 、 μ 固定,等待队列 $E(L_q)$ 会不断增大。由于此时 $\eta=0.3$ 匹配成功速率较低,尽管 λ 变化较大,单位时间内匹配成功的顾客数量相对来说还是较少,从而导致 $E(C_s)$ 、 $E(B_s)$ 整体变化不大。这表明当 λ 增大时,若此时系统不采取激励措施来提高匹配成功速率,会造成顾客在等待队列中出现不耐烦的顾客行为进而离开系统,对系统收益造成一定的损失。

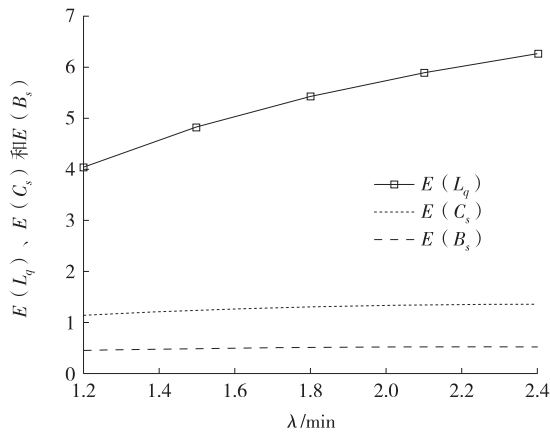


图 1 λ 对 $E(L_q)$ 、 $E(C_s)$ 和 $E(B_s)$ 的影响变化

Fig. 1 Change of influence of λ on $E(L_q)$, $E(C_s)$ and $E(B_s)$

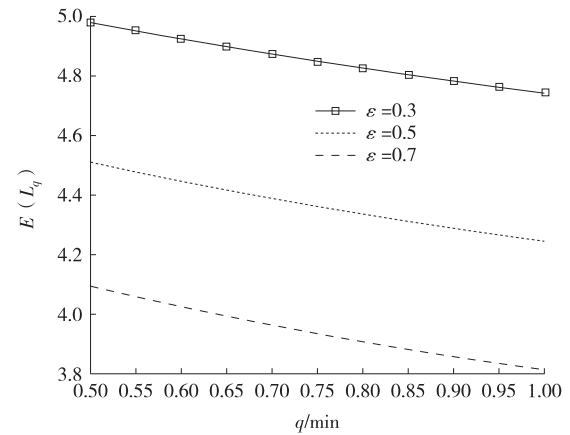


图 2 μ 、 η 对 $E(L_q)$ 的影响

Fig. 2 Change of influence of μ, η on $E(L_q)$

当 $\lambda=1.5, q=1, \eta=0.3$ 时,图2反映了 $E(L_q)$ 随 μ 、 ϵ 的变化趋势。不难发现当 ϵ 不变时, $E(L_q)$ 减少的相对缓慢,这是由于顾客在接受服务前,需要与服务台进行接触匹配,当其他系统参数不变时, μ 的增大对于 $E(L_q)$ 的影响不会很大。而当服务台拒绝匹配速率 ϵ 增大时, $E(L_q)$ 则下降十分明显。这表明当系统无法采用激励措施提高匹配成功速率时,由于 $E(L_q)$ 较大时会给系统带来不必要的成本,系统会为了减少成本而提高服务台拒绝匹配速率 ϵ ,来使得系统运营成本较低。

通过上述分析可知,当其他系统参数不变时,仅有服务率 μ 的变化,对于系统性能指标影响较小。下面分别讨论当 μ 与 ϵ 、 η 共同变化时对 $E(C_s)$ 、 $E(B_s)$ 的影响。

当 $\lambda=1.5, \epsilon=0.3, q=1$ 时,表1反映了 μ 、 η 对 $E(C_s)$ 、 $E(B_s)$ 的影响。当 μ 、 η 同时增大时, $E(C_s)$ 在不断减小, $E(B_s)$ 在不断增大。这表明了随着 η 的增大,匹配成功的顾客数量越来越多,尽管 μ 的增大会使得服务时间变短,但由于顾客数量的增大,处于忙期的服务台数量仍然会小幅度增大。当 $\lambda=1.5, \eta=0.3, q=1$ 时,表2反映了 μ 、 ϵ 对 $E(C_s)$ 、 $E(B_s)$ 的影响。当 μ 、 ϵ 同时增大时, $E(C_s)$ 、 $E(B_s)$ 均不断减小。一方面 ϵ 的增大使得服务台拒绝匹配的速率增大,会使得继续接受服务的顾客数量越来越少,且随着 μ 的增大会使得服务时间变短,此时 $E(B_s)$ 下降明显。

表 1 系统参数 μ 、 η 对 $E(C_s)$ 、 $E(B_s)$ 的影响

Tab. 1 Effects of system parameters μ, η on $E(C_s), E(B_s)$

系统参数	$E(C_s)$	$E(B_s)$
$\mu=0.5, \eta=0.3$	1.220 9	0.601 1
$\mu=0.6, \eta=0.4$	1.175 4	0.626 4
$\mu=0.7, \eta=0.5$	1.133 7	0.643 1
$\mu=0.8, \eta=0.6$	1.094 7	0.653 2
$\mu=0.9, \eta=0.7$	1.057 9	0.658 4
$\mu=1.0, \eta=0.8$	0.993 5	0.659 6

表 2 系统参数 μ 、 ϵ 对 $E(C_s)$ 、 $E(B_s)$ 的影响

Tab. 2 Effects of system parameters μ, ϵ on $E(C_s), E(B_s)$

系统参数	$E(C_s)$	$E(B_s)$
$\mu=0.5, \epsilon=0.3$	1.220 9	0.601 1
$\mu=0.6, \epsilon=0.4$	1.181 7	0.533 2
$\mu=0.7, \epsilon=0.5$	1.138 3	0.471 7
$\mu=0.8, \epsilon=0.6$	1.091 7	0.416 3
$\mu=0.9, \epsilon=0.7$	1.043 1	0.366 9
$\mu=1.0, \epsilon=0.8$	1.000 1	0.323 3

当系统等待顾客较多时,若系统的运营能力较强,可以同时增大 μ 、 η 使得更多的顾客接触匹配成功的同时,服务时间也会变少,系统收益会增大。与此相反的是,若由于特殊情况比如极端天气、极端路况导致系统的运营能力会下降,系统会提高 μ 、 ϵ ,使得等待队列顾客数量下降,系统成本下降,从而减少系统损失。

由于排队系统的复杂性,下面从服务台角度出发,研究服务台进入系统速率 q 对系统性能指标的影响。当 $\eta=0.3$, $\mu=1$, $\epsilon=0.3$ 时,图 3 反映了 $E(L_q)$ 随 q 、 λ 的变化趋势。当 q 不变, λ 增大时, $E(L_q)$ 会变大。当 λ 不变,随着 q 的增大, $E(L_q)$ 会不断变小。这是由于当系统中的服务台都处于忙期或接触匹配期时,新的服务台进入系统的速率增大, $E(L_q)$ 会减少。即当系统中 $E(L_q)$ 较大,系统会采取激励措施使得服务台进入系统的速率增大,从而使得 $E(L_q)$ 减少,进入系统接受服务的顾客数量更多,系统收益更大。

当 $\lambda=3$, $\mu=1$, $\epsilon=0.3$ 时,图 4 反映了 $E(B_s)$ 随 q 、 η 的变化趋势。当 q 不变,由于 η 的增大使得系统接触匹配成功率增大,顾客接触匹配成功的顾客数量会越多,服务台处于忙期的数量会越来越多。同样的,当 η 不变, q 的增大会使得 $E(B_s)$ 增大,这是因为服务台进入系统的速率增大,系统中会有更多的服务台与顾客进行接触匹配,顾客接触匹配成功的顾客数量越多,服务台处于忙期的数量会越来越多。

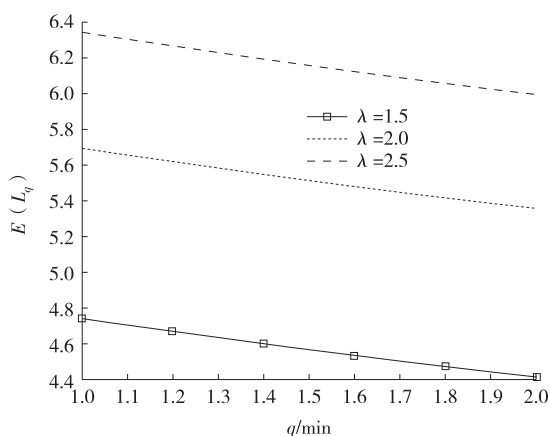


图 3 q 、 λ 对 $E(L_q)$ 的影响

Fig. 3 Change of influence of q , λ on $E(L_q)$

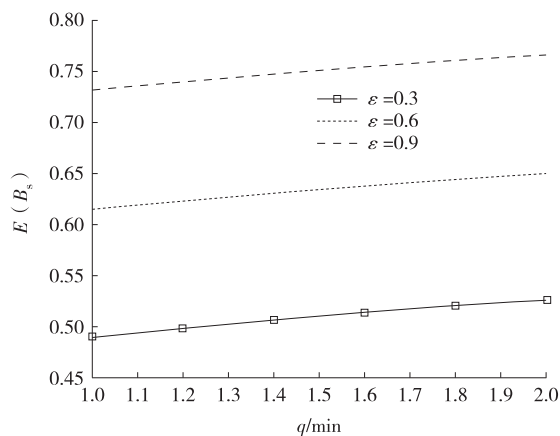


图 4 q 、 η 对 $E(L_q)$ 的影响变化

Fig. 4 Change of influence of q , η on $E(L_q)$

4.2 系统收益分析

为了进一步分析系统收益的合理性,下面通过构造单位时间内预期收益函数来分析排队系统。现实生活中若顾客等待时间较长,顾客在下次接受服务的意愿就会大大减小。为此,定义 c_w 为等待队列每位顾客单位时间对系统所造成的损失收益, c_k 为正在接受服务的每位顾客单位时间对系统带来的收益, c_f 为单位时间每个服务台处于接触匹配期所带来的潜在收益, c_r 为单位时间内随服务率变化的处于忙期状态的服务成本,则单位时间系统的收益函数为:

$$T = c_k E(B_s) - c_r E(B_s) - c_f E(C_s) - c_w E(L_q)。$$

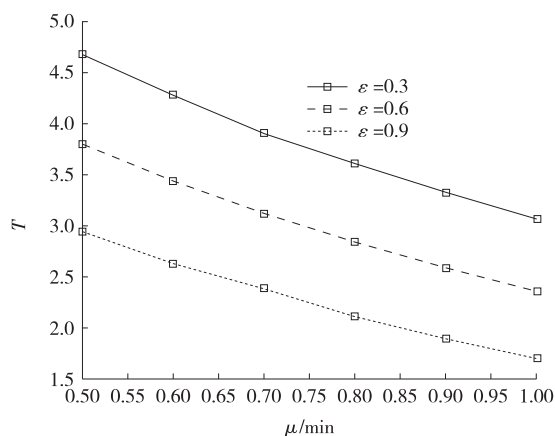
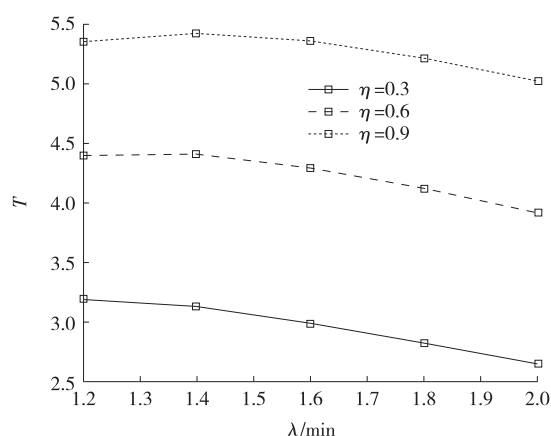
为保证系统盈利,需满足:

$$c_k > \frac{c_r E(B_s) + c_f E(C_s) + c_w E(L_q)}{E(B_s)}。$$

为更加方便直观地分析不同变量对收益函数的影响,假设 $c_w=1$, $c_f=2$, $c_r=3$, $c_k=15$, $c=3$, $N=13$, $q=1$, $b=0.5$, $\sigma=0.3$,研究顾客到达率 λ ,服务率 μ ,接触匹配率 η ,服务台拒绝匹配率 ϵ 对系统收益的影响,见图 5 和图 6。

当 $\mu=1$, $\epsilon=0.3$ 时,图 5 反映了 T 随 λ 、 η 的变化趋势。当 λ 不变时,随着 η 的增大,匹配成功的顾客越来越多,系统收益也会越来越大。当 η 不变且较小时,由于系统接触匹配成功的顾客数量较少,系统接受顾客进入系统的能力有限,随着 λ 的增大导致 $E(L_q)$ 较大使得系统成本增大,导致系统收益降低。当 $\eta=0.9$ 较大时, T 随着 λ 增大先增大后减小。这是由于 η 增大使得系统接触匹配成功的顾客数量更多,系统有能力接受更多的顾客进入系统使得系统收益增大,但当 λ 越来越大时,系统接受顾客数量的能力达到阈值,则系统成本增大,系统收益减小。

当 $\lambda=1.5, \eta=0.3$ 时,图6反映了 T 随 μ, ϵ 的变化趋势。当 ϵ 不变时, T 随着 μ 的增大而减小,这是由于此时 λ, η 较小, μ 的增大使得系统成本增大,但所带来的系统收益却较小,使得 T 下降,这表明当进入系统的顾客数量较少时,系统需要选择低成本的服务速率来使得系统收益最大。当 μ 不变时, T 随着 ϵ 增大而减小。由于 ϵ 的增大导致服务台拒绝匹配的顾客数量越来越多,给系统收益带来了一定损失,使得系统收益较低,系统也应当对服务台采取激励措施,比如每单的系统抽成比适当降低来提高系统中服务台的积极性,使得系统收益更大。

图5 λ, η 对 T 的影响变化Fig. 5 Change of influence of λ, η on T 图6 μ, ϵ 对 T 的影响变化Fig. 6 Change of influence of μ, ϵ on T

5 结论

本文从手机打车软件的实际排队场景出发,在 M/M/c 排队模型的基础上加入了可变到达率、可变服务率、接触匹配等策略,更加贴合实际情况。本文中服务台进入系统与传统服务台休假策略不同的是,服务台更多的是考虑自己的个人收益而主动加入系统进行服务,这也与生活中许多网络线上平台的模式相同。本文通过数值例子详细评估了系统参数的变化对系统性能指标及系统收益的影响。由于系统处于完全可见情形,等待队列的变化会导致顾客到达率的减小,系统考虑系统收益会采取提高服务率等策略来使得系统收益最大。系统根据等待队列大小来动态调整服务率等性能指标的同时,也需要采取一定的激励措施来提高接触匹配成功率等性能指标。激励措施可以是双向的,在顾客到达率较高时,可以降低对服务台的抽成比激励更多的服务台进入系统。在顾客到达率较低时,可以降低订单价格来激励更多顾客进入系统且提高了接触匹配率,以此来提高系统收益。

参考文献:

- [1] HAIGHT F A. Queueing with reneging[J]. Metrika, 1959, 2(1): 186-197.
- [2] 田乃硕. 控制休假的 M/M/1 排队模型及其优化[J]. 系统工程, 1989(2): 40-46.
TIAN N S. A queue M/M/1 with controllable serve's vacation and its optimization[J]. Systems Engineering, 1989(2): 40-46.
- [3] BAI J R, SO K C, TANG C S, et al. Coordinating supply and demand on an on-demand service platform with impatient customers [J]. Social Science Electronic Publishing, 2019, 21(3): 556-570.
- [4] SUDTACHAT K, TANTRAI R A T N S, PHUNPENG V. The queueing model for perishable inventory with lost sale under random demand, lead time and lifetime[EB/OL]. (2017-01-20) [2022-05-19]. <https://www.actapress.com/PaperInfo.aspx?paperId=456370>.
- [5] ASHOUR M, LE-NGOC T. Priority queueing of long-range dependent traffic[J]. IEEE Computer Communications, 2008, 31, 3954-3963.
- [6] 卢宜玲, 李军祥, 齐缘. 疫情下考虑疲劳度的可变服务率排队模型[J]. 上海理工大学学报, 2021, 44(3): 288-298.
LU Y L, LI J X, QI Y. Queueing model with variable service rate considering fatigue under epidemic situation[J]. University of Shanghai for Science and Technology, 2021, 44(3): 288-298.
- [7] WANG Q, ZHANG B. Analysis of a busy period queueing system with balking, reneging and motivating[J]. Applied Mathematical Modelling, 2018, 64: 480-488.

- [8] KE J C, CHANG F M, LIU T H. M/M/c balking retrial queue with vacation[J]. Quality Technology and Quantitative Management, 2019, 16(1): 54-66.
- [9] 孙敬焄, 高郡, 张峻川, 等. 带接触匹配的 M/M/c 排队模型[J]. 数学的实践与认识, 2020, 50(19): 186-192.
SUN J X, GAO J, ZHANG J C, et al. M/M/c Queueing model with the state of matching and abandoning[J]. Mathematics in Practice and Theory, 2020, 50(19): 186-192.
- [10] GEORGE D K, XIA C H. Fleet-sizing and service availability for a vehicle rental system via closed queueing networks[J]. European Journal of Operational Research, 2011, 211(1): 198-207.
- [11] SHI Y, LIAN Z T, SHANG W X. Study of a passenger-taxi queueing system with nonzero matching time[C]//2015 12th International Conference on Service Systems and Service Management. Guangzhou: IEEE, 2015: 15344203.
- [12] QIAN X W, ZHANG W B, UKKUSURI S V, et al. Optimal assignment and incentive design in the taxi group ride problem[J]. Transportation Research Part B: Methodological, 2017, 103: 208-226.
- [13] ZHANG W, HONNAPPA H, UKKUSURI S V. Modeling urban taxi services with e-hailings: a queueing network approach [J]. Transportation Research Procedia, 2019, 38: 751-771.
- [14] NEUTS M F. Matrix-geometric solutions in stochastic models[M]. Baltimore: Johns Hopkins University Press, 1981.
- [15] 田乃硕, 岳德权. 拟生灭过程与矩阵几何解[M]. 北京: 科学出版社, 2002.
TIAN N S, YUE D Q. The quasi birth and death process and matrix-geometric solution[M]. Beijing: Science Press, 2002.

Operations Research and Cybernetics

Analysis of Contact Matching Model Based on M/M/c Queue with Variable Rate

TAN Yuting, XU Xiuli

(School of Science, Yanshan University, Qinhuangdao Hebei 066004, China)

Abstract: [Purposes] Combined with the actual queueing scene of DiDi taxi in life, it studies the M/M/c queueing model of contact matching with variable arrival rate and variable service rate. [Methods] It obtains the state transition law of quasi birth and death process and the infinitesimal generator matrix with probability analysis method, and gives the steady-state condition, steady-state probability distribution and main performance indices of the system by using the matrix geometric solution method. [Findings] The numerical analysis is presented to demonstrate the influence of system parameters on performance indices. The system profit function is established to discuss system optimal profit. [Conclusions] The influence relationship of system parameters on system performance indices and system profit are obtained, which shows the service rate and contact matching success rate in contact matching model should be adjusted at the same time in order to significantly improve the system benefit.

Keywords: queueing model; variable arrival rate; variable service rate; contact matching

(责任编辑 许 甲)