

具有平行顾客的双端排队系统的均衡分析

李远鹏, 丁栋山

燕山大学理学院统计系, 河北 秦皇岛

收稿日期: 2025年12月27日; 录用日期: 2026年1月23日; 发布日期: 2026年2月2日

摘要

本文研究了一类具有两类平行顾客的经典双端排队系统的出租车 - 乘客匹配过程。在经典模型中, 顾客采用个人效用最大化的均衡阈值策略和社会收益最大化的最优阈值策略加入系统, 推导了系统的稳态概率、性能指标及社会收益函数, 并通过数值实验探讨了相关参数对均衡阈值和社会最优阈值的影响以及对社会收益的影响。

关键词

双端排队, 平行顾客, 阈值策略, 社会收益

Equilibrium Analysis of a Double-Ended Queueing System with Parallel Customers

Yuanpeng Li, Dongshan Ding

Department of Statistics, School of Science, Yanshan University, Qinhuangdao Hebei

Received: December 27, 2025; accepted: January 23, 2026; published: February 2, 2026

Abstract

This paper studies the taxi-passenger matching process in a class of classical two-sided queueing systems with two types of parallel customers. In the classical model, customers join the system using an equilibrium threshold strategy that maximizes personal utility and an optimal threshold strategy that maximizes social welfare. The steady-state probabilities, performance indicators, and social welfare function of the system are derived. Numerical experiments are conducted to explore the effects of relevant parameters on the equilibrium threshold, the socially optimal threshold, and their impact on social welfare.

Keywords

Double-Ended Queueing, Parallel Customers, Threshold Strategy, Social Welfare

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

排队是生活中常见的一种现象, 体现在早晚高峰排队乘车、排队购票等。排队论则是研究随机排队服务系统的随机现象和工作过程的一种数学理论和方法, 在许多领域得到了广泛的应用。所谓的双端排队系统本质上是具有双边输入的匹配排队系统, 即来自顾客端和服务器端的双到达流。随着共享经济与服务系统的快速发展, 双端排队系统作为一种描述供需匹配过程的重要建模工具, 在交通出行、共享资源调度、餐饮服务等多个领域展现出广泛的应用价值。本文研究的“平行顾客”是指在同一服务系统中并行存在、但具有不同经济参数(如服务报酬、等待成本)的两类顾客。他们共享同一服务台(出租车), 遵循相同的服务规则(如 FCFS), 但基于自身不同的收益 - 成本结构做出独立的加入或退出决策。这与传统的“多等级服务”模型存在本质区别: 在多等级服务系统中, 不同类别顾客通常具有明确的服务优先级, 高优先级顾客可中断或抢占低优先级顾客的服务; 而本文中的平行顾客之间并无服务优先权差异, 其类别仅体现在决策参数的不同。此外, 本文模型亦不同于“重试排队”模型: 在重试排队中, 受阻顾客会离开服务区并在随机时间后重新尝试, 系统动态受重试过程影响; 而本文假设顾客到达后即做出不可撤销的加入或止步决策, 系统动态完全由初始到达与阈值策略决定。

引入第二类顾客后, 系统动力学的核心变化体现在: 策略交互与阈值耦合: 系统不再仅由单一顾客群体的决策阈值决定, 而是由两类顾客的阈值共同刻画其行为。这种耦合决策机制是单类顾客模型所不具备的; 状态空间的非对称分割: 系统的状态转移过程被两类顾客的阈值 n_1 和 n_2 分割为三个具有不同动态特征的区域, 这种由差异化经济参数驱动的、非对称的状态依赖行为, 使得系统动态比单类阈值模型更为复杂, 也更能反映现实世界中异质用户群体的真实决策模式; 社会收益函数的多元化结构: 社会收益函数需要同时权衡两类异质顾客群体的净收益与出租车运营方的收益。优化社会收益的目标, 实质上是在两类顾客的加入激励(创造匹配价值)与由此引发的拥堵成本(增加等待时间)之间寻求平衡, 同时还要考虑出租车资源的利用效率。

双端乘客 - 出租车排队问题最早是由 Kendall [1] 提出的, 他引入了一个关于乘客和出租车的例子: 在一个有序的出租车队列中, 一边是等待出租车的一群乘客的到来而形成的队列, 另一边是等待乘客的一群出租车。Dobbie [2] 发现了乘客和出租车非均匀泊松到达下的暂态行为。Kashyap [3] 考虑了出租车和乘客等待空间有限的双端队列, 使用补充变量技术来讨论平均出租车和一般和泊松到达过程中有限缓冲下的顾客队列长度。Manou 等[4] 考虑了一种运输设施, 其中顾客根据泊松过程到达, 该设施根据其容量为随机数量的顾客提供服务。他们得出了两种信息水平下的顾客策略。然而, 他们没有考虑缓冲区大小的重要作用。Shi 等[5] 研究了出租车等待空间有限的乘客 - 出租车问题, 研究了可观察和不可观察两种情况下乘客的策略行为。此外, 作者还分析了出租车缓冲规模对社会福利的影响, 得出了最优出租车缓冲规模。Shi 等[6] 也讨论了乘客 - 出租车问题, 乘客和出租车的等待空间都是有限的。Qihui Bu 等[7] 提供了对双端队列系统在随机服务能力下和犹豫客户行为影响下的静态分析的深入理解, 以及对管理层如何

制定信息披露政策的指导。平行顾客通常指具有不同服务需求(成本结构不同)的两类顾客, 在排队系统中结合重试策略和两类平行顾客, 在满足多样化客户需求的同时, 最大化系统的运营效率和经济效益。Liu 等[8]研究了一个不可靠的 $M/M/1 + 1$ 重试排队系统, 该系统具有两类并行顾客和一个备用服务器, 备用服务器在系统故障时提供批量服务。研究重点是从经济视角分析顾客的均衡策略和社会效益。张淞钦等[9]研究了一个具有两类平行顾客、服务器不完全故障且带延迟维修特性的排队系统, 在完全可见和几乎可见两种信息水平下重点分析了顾客的均衡策略行为和社会效益。徐秀丽等[10]研究了一个具有两类平行顾客和灾难清空的 $M/M/1$ 排队系统。在完全可见和几乎可见两种信息情形下, 分析了顾客的均衡阈值策略和系统的平均社会收益。

本文其余部分的内容如下。在第 2 节中, 对于具有两类平行顾客的经典双端排队模型进行了详细的描述, 并给出了平稳方程。在 3 节中, 研究了由平稳方程推出的稳态分布以及性能指标。在第 4 节中, 分析了顾客进入系统的策略问题。在第 5 节中, 结合数据进行数值分析。最后在第 6 节中, 总结全文并给出可扩展的研究方向。

2. 模型建立

2.1. 模型描述

在本节中, 我们考虑了一个乘客 - 出租车服务系统, 它是出租车和乘客之间的匹配过程的一个双端排队系统, 现在我们给出该模型的详细描述。我们假设有两类平行顾客, 分别用 1 类顾客和 2 类顾客来表示, 两类顾客的到达均遵循泊松过程, 到达率分别是 λ_1 和 λ_2 , 因此顾客总到达速率是 $\lambda = \lambda_1 + \lambda_2$ 。我们假设出租车也按照泊松过程到达, 当系统中的乘客数量大于 0 时, 到达率设为 μ_2 , 否则为 μ_1 , 且有 $\mu_2 > \mu_1$ 。在出租车站, 出租车的容量是 N , 这意味着如果有 N 出租车在等乘客, 后续的出租车就不能加入出租车站。乘客可以不受限制地进入出租车站。设 $N(t)$ 表示出租车站的乘客或出租车在 t 时刻的排队长度。

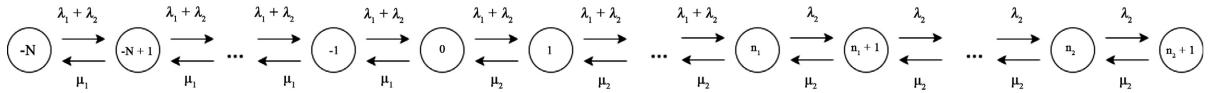
若 $N(t) > 0$ 则表示有乘客在等出租车; $N(t) = 0$ 则表示系统是空闲的; $N(t) < 0$ 则表示有出租车正在等待新乘客的到来。我们假设顾客遵循先到先服务的规则, 如果出租车载有一名乘客, 立即离开出租车站, 因此假设出租车和乘客的匹配时间为零。此外, 我们假设两类顾客和出租车的到达时间间隔、匹配时间是相互独立的。

假设每一个完成服务的 j 类顾客获得 R_j ($j = 1, 2$) 单位的奖励, 同时每一个加入系统的 j 类顾客在单位时间内产生与服务相关的等待成本是 C_j ($j = 1, 2$), 出租车的单位时间等待成本是 C_0 , 且出租车每完成一次服务的收入是 p 即乘客支付的费用。因此本节的双端排队模型可以表示为一维排队过程, $\{N(t), t \geq 0\}$ 是一个马尔可夫过程, 其状态空间为 $\Omega = \{-N, -N+1, \dots, -1, 0, 1, \dots\}$ 。设状态 i 的稳态概率为 π_i , 则稳态概率可以表示为

$$\pi_i = \lim_{t \rightarrow \infty} P\{N(t) = i\}, i \in \Omega \quad (1)$$

在完全可见的情况下, 到达的顾客知道系统中的队列长度和出租车数量, 同时也知道自身所属的顾客类型。假定第 j 类顾客进入系统排队的进入阈值是 n_j ($j = 1, 2$), 即当系统中的某类顾客数小于自身进入阈值时, 则这类顾客选择进入; 否则止步。两类顾客的阈值顺序将由他们的净收益结构决定, 即取决于 $\frac{R_j - p}{C_j}$ 的相对大小。若 $\frac{R_1 - p}{C_1} > \frac{R_2 - p}{C_2}$, 则第 1 类顾客更有耐心等待, 可能出现 $n_1 \geq n_2$; 反之则 $n_2 \geq n_1$ 。

不妨假设第一类顾客先止步, 即有 $n_2 > n_1$ 。状态转移图图 1 如下:

**Figure 1.** State transition diagram**图 1.** 状态转移图

2.2. 平稳方程

根据上述状态转移图, 得到如下平衡方程,

$$(\lambda_1 + \lambda_2)\pi_{-N} = \mu_1\pi_{-N} \quad (2)$$

$$(\lambda_1 + \lambda_2 + \mu_1)\pi_i = (\lambda_1 + \lambda_2)\pi_{i+1} + \mu_1\pi_{i-1}, N < i < 0 \quad (3)$$

$$(\lambda_1 + \lambda_2 + \mu_2)\pi_i = (\lambda_1 + \lambda_2)\pi_{i-1} + \mu_2\pi_{i+1}, 0 \leq i < n_1 \quad (4)$$

$$(\lambda_2 + \mu_2)\pi_i = \lambda_2\pi_{i-1} + \mu_2\pi_{i+1}, n_1 \leq i < n_2 \quad (5)$$

$$\mu_2\pi_2 = \lambda_2\pi_{n_2-1} \quad (6)$$

3. 稳态分析

3.1. 平稳分布

在完全可观测情况下的平稳分布如下:

$$\pi_i = \pi_{-N}\rho_1^{N+i}, -N < i \leq 0 \quad (7)$$

$$\pi_i = \pi_0\rho_2^i, 0 < i < n_1 \quad (8)$$

$$\pi_i = \pi_{n_1}\rho_3^{i-n_1}, n_1 \leq i \leq n_2 \quad (9)$$

$$\pi_{-N} = \frac{(1-\rho_1)(1-\rho_2)(1-\rho_3)}{(1-\rho_1^{N+1})(1-\rho_2^{n_1})(1-\rho_3^{n_2-n_1}) + \rho_1^N\rho_2^{n_1}(1-\rho_3^{n_2-n_1})(1-\rho_1)} \quad (10)$$

其中 $\rho_1 = \frac{\lambda_1 + \lambda_2}{\mu_1}$, $\rho_2 = \frac{\lambda_1 + \lambda_2}{\mu_2}$, $\rho_3 = \frac{\lambda_2}{\mu_2}$ 。

经过代数计算和推导, 得到 π_i , 并由下述归一化条件得到 π_{-N} 。

$$\pi_{-N} \frac{1-\rho_1^{N+1}}{1-\rho_1} + \pi_0 \frac{1-\rho_2^{n_1}}{1-\rho_2} + \pi_{n_1} \frac{1-\rho_3^{n_2-n_1}}{1-\rho_3} = 1 \quad (11)$$

3.2. 性能指标

对于完全可见的情况, $E(L_1)$ 和 $E(L_2)$ 分别表示为 1 类顾客和 2 类顾客的预期队列长度, $E(L_t)$ 表示出租车的预期队列长度, 其形式如下:

$$E(L_1) = \sum_{i=1}^{n_1} i\pi_i = \pi_0 \frac{\rho_2 \left[1 - (n_1 + 1)\rho_2^{n_1} + n_1\rho_2^{n_1+1} \right]}{(1-\rho_2)^2} \quad (12)$$

$$E(L_2) = \sum_{i=1}^{n_2} i\pi_i = \frac{\lambda_2}{\lambda_1 + \lambda_2} \pi_0 \frac{\rho_2 \left[1 - (n_1 + 1)\rho_2^{n_1} + n_1\rho_2^{n_1+1} \right]}{(1-\rho_2)^2} + \pi_{n_1} \frac{\rho_3 \left[1 - (n_2 - n_1 + 1)\rho_3^{n_2-n_1} + (n_2 - n_1)\rho_3^{n_2-n_1+1} \right]}{(1-\rho_3)^2} \quad (13)$$

$$E(L_t) = \sum_{i=-N}^0 (-i)\pi_i = \pi_{-N} \frac{N\rho_1^{N+1} - (N+1)\rho_1^N + \rho_1}{(1-\rho_1)^2} \quad (14)$$

证明: 根据定义 $E(L_1) = \sum_{i=1}^{n_1} i \cdot \pi_i$ 。

对于 $1 \leq i \leq n_1$, 系统中两类顾客均存在, 则有 $\pi_i = \pi_0 \rho_2^i$, 代入上式有 $E(L_1) = \pi_0 \sum_{i=1}^{n_1} i \cdot \rho_2^i$ 。

利用几何级数求和公式: $\sum_{i=1}^n i r^i = \frac{r[1 - (n+1)r^n + nr^{n+1}]}{(1-r)^2}, |r| < 1$

令 $r = \rho_2$, $n = n_1$ 得:

$$E(L_1) = \pi_0 \frac{\rho_2 [1 - (n_1+1)\rho_2^{n_1} + n_1\rho_2^{n_1+1}]}{(1-\rho_2)^2} \quad (15)$$

$E(L_1) = \sum_{i=1}^{n_1} i \cdot \pi_i$ 这是所有顾客(1类和2类)的期望队列长度在 $i \leq n_1$ 的部分。实际上, 在 $i \leq n_1$ 时, 两类顾客比例与到达率成比例, 因此2类顾客在这部分的期望队列长度为:

$$\frac{\lambda_2}{\lambda_1 + \lambda_2} \pi_0 \sum_{i=1}^{n_1} i \rho_2^i \quad (16)$$

当 $n_1 \leq i \leq n_2$ 时: $\pi_i = \pi_{n_1} \rho_3^{i-n_1}$ 。

于是:

$$\sum_{i=n_1}^{n_2} i \pi_i = \pi_{n_1} \sum_{k=0}^{n_2-n_1} k \rho_3^k \quad (17)$$

利用几何级数求和有:

$$\sum_{i=n_1}^{n_2} i \pi_i = \pi_{n_1} \frac{\rho_3 [1 - (n_2 - n_1 + 1)\rho_3^{n_2-n_1} + (n_2 - n_1)\rho_3^{n_2-n_1+1}]}{(1-\rho_3)^2} \quad (18)$$

因此

$$E(L_2) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \pi_0 \frac{\rho_2 [1 - (n_1+1)\rho_2^{n_1} + n_1\rho_2^{n_1+1}]}{(1-\rho_2)^2} + \pi_{n_1} \frac{\rho_3 [1 - (n_2 - n_1 + 1)\rho_3^{n_2-n_1} + (n_2 - n_1)\rho_3^{n_2-n_1+1}]}{(1-\rho_3)^2} \quad (19)$$

证毕。

对于完全可见的情况, $E(W_1)$ 和 $E(W_2)$ 分别表示为 1类顾客和 2类顾客的平均等待时间, $E(W_t)$ 表示出租车的平均等待时间, 其形式如下:

$$E(W_1) = \frac{E(L_1)}{\lambda_1} = \frac{\pi_0}{\lambda_1} \frac{\rho_2 [1 - (n_1+1)\rho_2^{n_1} + n_1\rho_2^{n_1+1}]}{(1-\rho_2)^2} \quad (20)$$

$$E(W_2) = \frac{E(L_2)}{\lambda_2} = \frac{\pi_0}{\lambda_2} \frac{\rho_2 [1 - (n_1+1)\rho_2^{n_1} + n_1\rho_2^{n_1+1}]}{(1-\rho_2)^2} + \frac{\pi_{n_1}}{\lambda_2} \frac{\rho_3 [1 - (n_2 - n_1 + 1)\rho_3^{n_2-n_1} + (n_2 - n_1)\rho_3^{n_2-n_1+1}]}{(1-\rho_3)^2} \quad (21)$$

$$E(W_t) = \frac{E(W_t)}{\mu_{eff}} = \pi_{-N} \frac{N\rho_1^{N+1} - (N+1)\rho_1^N + \rho_1}{\mu_{eff} (1-\rho_1)^2} \quad (22)$$

μ_{eff} 是出租车总的有效到达率, 其表达式是:

$$\mu_{eff} = \mu_1 \sum_{i=-N}^0 \pi_i + \mu_2 \sum_{i=1}^{n_2} \pi_i = \mu_1 \sum_{i=-N}^0 \pi_{-N} \rho_1^{N+i} + \mu_2 \left(\sum_{i=1}^{n_1-1} \pi_0 \rho_2^i + \sum_{i=n_1}^{n_2} \pi_{n_1} \rho_3^{i-n_1} \right) \quad (23)$$

在完全可见的 $M/M/1$ 经典双端排队系统中, 假设两类顾客遵循 (n_1, n_2) 的阈值策略, 则社会收益函数 $S(n_1, n_2)$ 为:

$$S(n_1, n_2) = \lambda_1 (R_1 - p - C_1 E(W_1)) + \lambda_2 (R_2 - p - C_2 E(W_2)) + \mu_{\text{eff}} (p - C_0 E(W_t)) \quad (24)$$

4. 阈值策略

4.1. 均衡阈值策略

对于完全可见的情况, 我们给出了两类顾客在到达后选择加入动态匹配的乘客 - 出租车双端队列系统的均衡策略如下:

$$n_j^e = \left\lfloor \frac{\mu_2 (R_j - p)}{C_j} - 1 \right\rfloor, j = 1, 2 \quad (25)$$

证明: 由两类顾客的效用函数 $U_j(n) = R_j - p - C_j \frac{n+1}{\mu_2}$, 因此两类顾客的阈值 n_j 需要满足以下不等式:

$$\begin{cases} U_j(n_j) \geq 0 \\ U_j(n_j + 1) < 0 \end{cases} \quad (26)$$

两类顾客的均衡阈值 n_j^e 均是正整数, 通过求解不等式可得到, 证毕。

4.2. 最优阈值策略

使社会收益函数 $S(n_1, n_2)$ 最大化时顾客的阈值策略 (n_1^*, n_2^*) 即是两类顾客的进入最优阈值。

5. 数值分析

5.1. 参数对等待时间的影响

我们首先研究了经典双端排队中的均衡策略下基本参数(两类顾客的到达率 λ_1 和 λ_2 以及出租车在有无乘客时的到达率 μ_1 和 μ_2)对两类顾客等待时间的影响。

图 2 展示了当固定 2 类顾客到达率是 $\lambda_2 = 0.5$ 时, λ_1 的变化对两类顾客等待时间的影响; 图 3 展示了当固定 1 类顾客到达率是 $\lambda_1 = 0.5$ 时, λ_1 的变化对两类顾客等待时间的影响($\mu_1 = 2$, $\mu_2 = 3$, $C_1 = 2$, $C_2 = 1$, $R_1 = 20$, $R_2 = 15$, $N = 4$)。

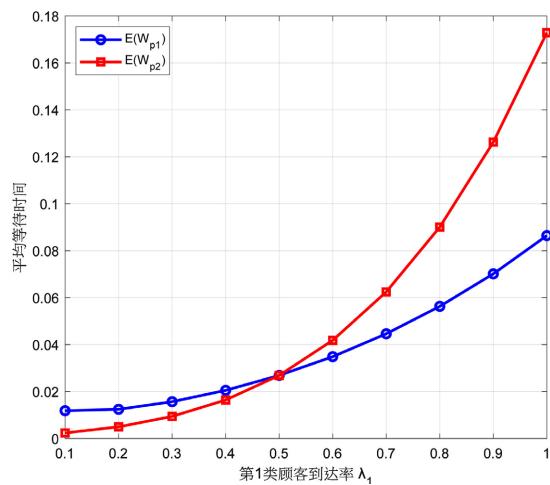


Figure 2. λ_1 impact on the average waiting time for customers

图 2. λ_1 对顾客的平均等待时间的影响

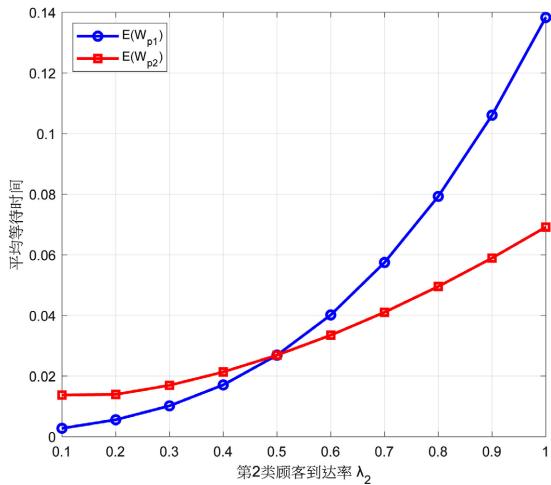


Figure 3. λ_2 impact on the average waiting time for customers
图 3. λ_2 对顾客的平均等待时间的影响

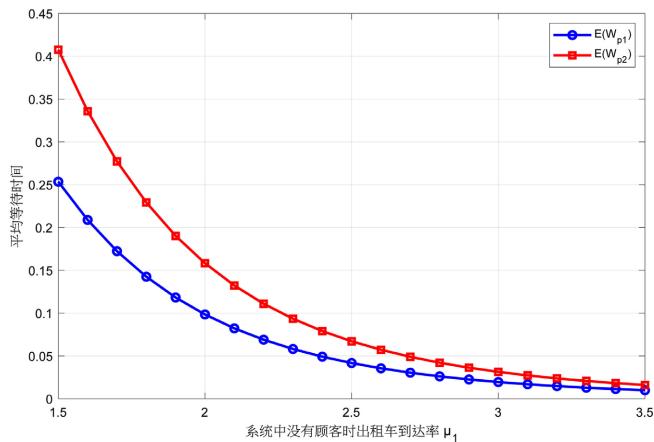


Figure 4. μ_1 impact on the average waiting time for customers
图 4. μ_1 对顾客的平均等待时间的影响

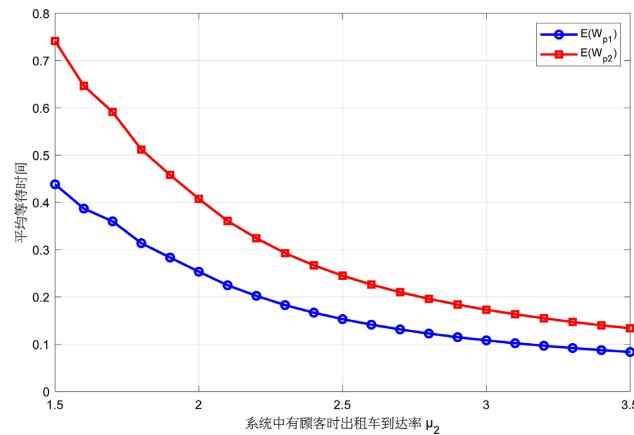


Figure 5. μ_2 impact on the average waiting time for customers
图 5. μ_2 对顾客的平均等待时间的影响

由图2可知, 随着 λ_1 从0.1增加到1, $E(W_1)$ 缓慢上升, 说明1类顾客的等待时间受自身到达率影响较小, $E(W_2)$ 随 λ_1 增加显著上升, 尤其是在 $\lambda_1 > 0.5$ 后增长更为明显, 说明2类顾客的等待时间对1类顾客到达率非常敏感; 图3反映了 $E(W_1)$ 在 λ_2 较低时保持平稳, 但随着 λ_2 接近1时开始显著上升, 说明在2类顾客的高到达率背景下, 1类顾客的等待时间也会受到较大影响。 $E(W_2)$ 整体高于图1中对应曲线, 且增长更为陡峭, 反映出在2类顾客高到达率下系统拥堵加剧, 1类顾客等待时间显著增加且急剧恶化。

图4展示了当固定系统有顾客时出租车到达率是 $\mu_2 = 2$ 时, μ_1 的变化对两类顾客等待时间的影响; 图5展示了当固定系统没有顾客时出租车到达率是 $\mu_1 = 1.5$ 时, μ_2 的变化对两类顾客等待时间的影响($\lambda_1 = 0.8$, $\lambda_2 = 0.5$, $C_1 = 2$, $C_2 = 1$, $R_1 = 15$, $R_2 = 12$, $N = 5$)。

图4体现了在 μ_1 较小的初始阶段, 两条曲线都处于高位, 且下降得非常迅速。这说明当空闲出租车稀少时, 稍微增加 μ_1 就能带来巨大的性能提升, 即等待时间对 μ_1 的敏感度非常高。随着 μ_1 继续增大, 两条曲线都逐渐变得平缓, 并趋近于零。这意味着当空闲出租车已经足够多时, 再继续增加其到达率, 对缩短等待时间的边际效益将逐渐递减。该图清晰地揭示了即使在繁忙状态 μ_2 下出租车到达率 μ_1 固定的前提下, 优化空闲状态出租车到达率 μ_1 也能极大改善顾客体验; 图5反映了在 μ_2 较小阶段(如1.5到2.5之间), 曲线下降尤为陡峭, 表明在此区域内提升繁忙期出租车资源, 对减少等待时间的效果非常显著。当 μ_2 超过3以后, 曲线逐渐平缓, 说明此时继续增加繁忙期的出租车到达率, 对进一步压缩等待时间的贡献逐渐减弱。上述两个图共同揭示了一个关键策略: 在资源有限的情况下, 同时优化系统在空闲状态和繁忙状态的服务资源分配, 是降低顾客等待时间的有效途径。

5.2. 顾客阈值策略的数值实验

图6展示了 R_1 对顾客均衡阈值和社会最优阈值的影响; 图7展示了 R_2 对顾客均衡阈值和社会最优阈值的影响($\mu_2 = 3$, $C_1 = 2$, $C_2 = 1$, $p = 5$)。

图6体现出 R_1 对2类顾客的均衡阈值没有影响, 对1类顾客的均衡阈值和两类顾客的社会最优阈值均是逐渐增加的, 1类顾客的均衡阈值增加较快, 社会最优阈值增加的较为缓慢, 但始终有 $n_2^* > n_1^*$; 图7反映出 R_2 对1类顾客的均衡阈值没有影响, 对2类顾客的均衡阈值和两类顾客的社会最优阈值均是逐渐增加的, 2类顾客的均衡阈值增加较快, 社会最优阈值增加的较为缓慢, 但始终有 $n_2^* > n_1^*$ 。

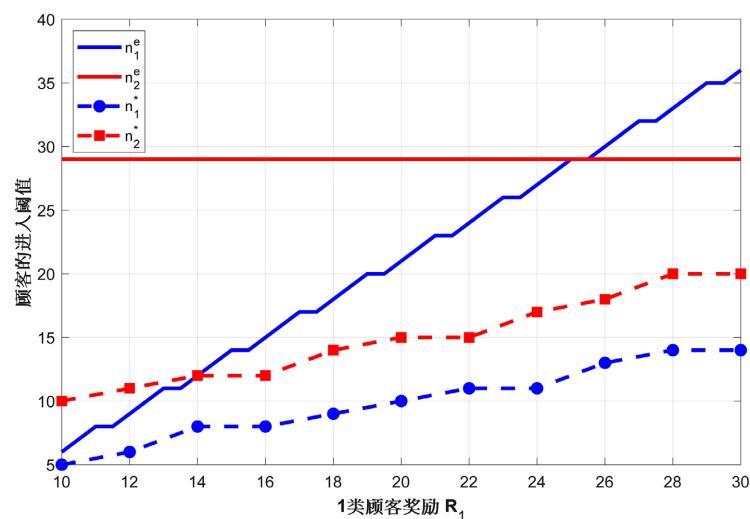


Figure 6. R_1 impact on customers' equilibrium threshold and the socially optimal threshold

图6. R_1 对顾客均衡阈值和社会最优阈值的影响

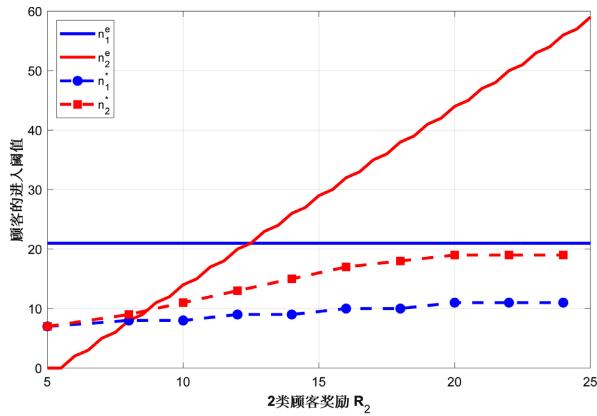


Figure 7. R_2 impact on customers' equilibrium threshold and the socially optimal threshold
图 7. R_2 对顾客均衡阈值和社会最优阈值的影响

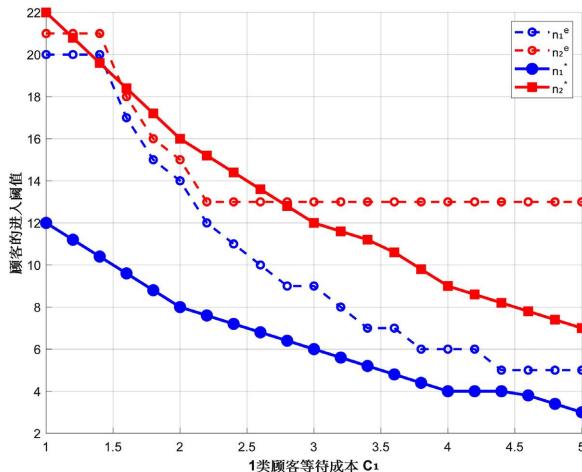


Figure 8. C_1 impact on customers' equilibrium threshold and the socially optimal threshold
图 8. C_1 对顾客均衡阈值和社会最优阈值的影响

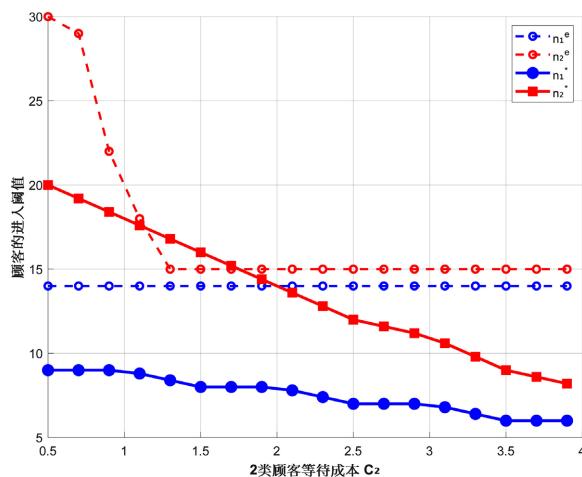


Figure 9. C_2 impact on customers' equilibrium threshold and the socially optimal threshold
图 9. C_2 对顾客均衡阈值和社会最优阈值的影响

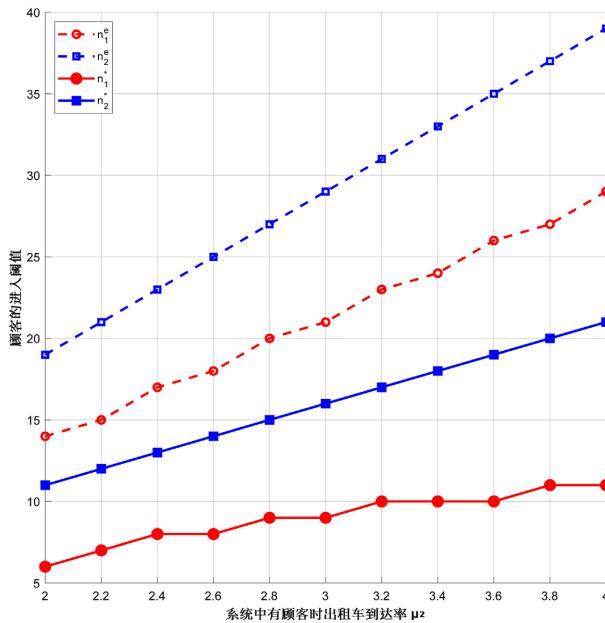
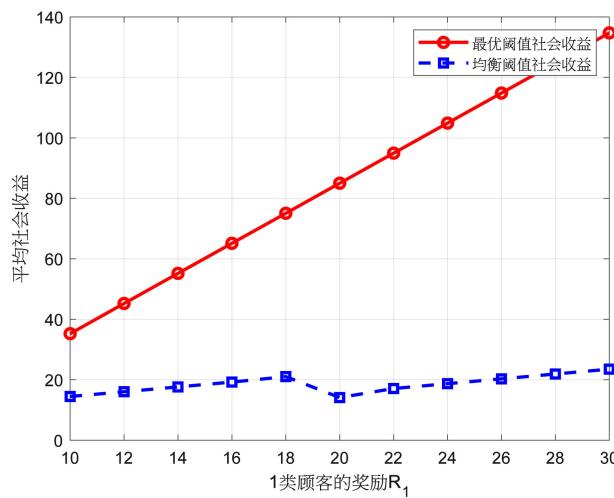
**Figure 10.** μ_2 impact on customers' equilibrium threshold and the socially optimal threshold**图 10.** μ_2 对顾客均衡阈值和社会最优阈值的影响

图 8 展示了 C_1 对顾客均衡阈值和社会最优阈值的影响; 图 9 展示了 C_2 对顾客均衡阈值和社会最优阈值的影响 ($\mu_2 = 3$, $R_1 = 15$, $R_2 = 12$, $p = 5$)。

图 8 和图 9 均体现了随着顾客成本的增加, 顾客的均衡阈值和社会最优阈值都是逐渐减少的, 对于 C_1 的影响, 顾客阈值间有 $n_2^* > n_2^e > n_1^e > n_1^*$ 这样的大小关系; 对于 C_2 的影响, 在 C_2 较小时, 顾客阈值间大小关系是 $n_2^e > n_2^* > n_1^e > n_1^*$, 在 C_2 较大时, 顾客阈值间大小关系是 $n_2^e > n_1^e > n_2^* > n_1^*$ 。

图 10 展示了 μ_2 对顾客均衡阈值和社会最优阈值的影响 ($R_1 = 15$, $R_2 = 12$, $C_1 = 2$, $C_2 = 1$, $p = 5$)。图 10 反映了随着出租车到达率的增加, 顾客的均衡阈值和社会最优阈值都是逐渐增加的, 且有 $n_2^e > n_1^e > n_2^* > n_1^*$ 。

**Figure 11.** R_1 impact on average social welfare**图 11.** R_1 对平均社会收益的影响

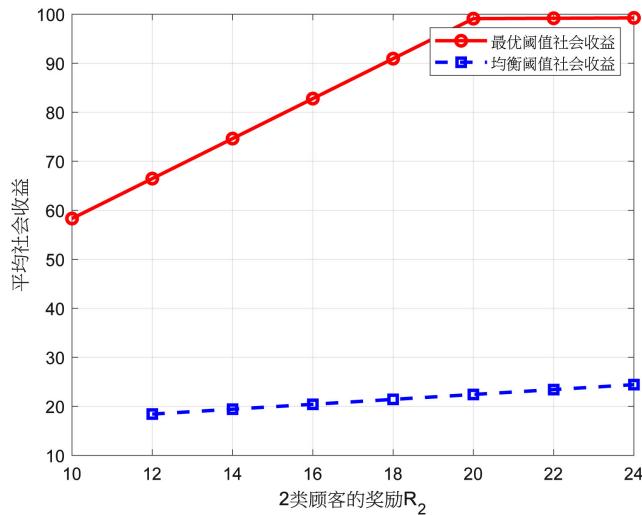
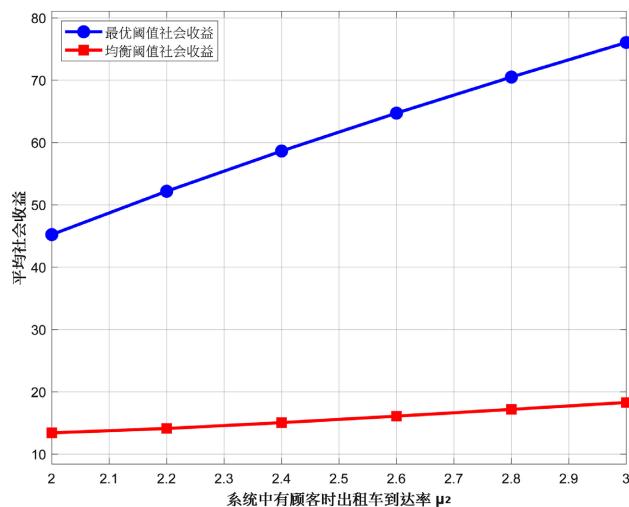
Figure 12. R_2 impact on average social welfare图 12. R_2 对平均社会收益的影响Figure 13. μ_2 impact on average social welfare图 13. μ_2 对平均社会收益的影响

图 11~图 13 描述了两类顾客分别在 R_1 , R_2 , μ_2 方面采取相应的均衡策略和社会最优策略时单位时间的社会效益。

5.3. 最优收费与补贴机制

在实际运营中, 乘客 - 出租车匹配系统往往存在“外部性”: 顾客在决定是否加入系统时仅考虑自身收益, 而忽略了其加入行为对其他顾客和出租车等待时间的影响, 从而导致系统整体效率下降(即社会收益未最大化)。为引导顾客行为趋近于社会最优, 运营商可通过定价机制(如动态收费或补贴)内部化这种外部性。

设运营商对每笔交易向乘客收取费用 p (固定部分), 并可额外施加一项“拥堵调节费” Δp_j (可正可负), 则第 p 类顾客的实际支付为:

$$p_j = p + \Delta p_j, j=1,2 \quad (27)$$

此时顾客的净收益函数变为:

$$U_j(n) = R_j - p + \Delta p_j - C_j \frac{n+1}{\mu_2} \quad (28)$$

由均衡阈值 n_j^e 需满足的不等式(26), 解得新的均衡阈值策略是:

$$n_j^e = \left\lfloor \frac{\mu_2(R_j - p - \Delta p_j)}{C_j} - 1 \right\rfloor \quad (29)$$

社会最优阈值 n_j^* 由最大化社会收益函数 $S(n_1, n_2)$ 得到, 为使 $n_j^e = n_j^*$, 应选择 Δp_j 满足:

$$\frac{\mu_2(R_j - p - \Delta p_j)}{C_j} - 1 = n_j^* \quad (30)$$

解得最优调节费是:

$$\Delta p_j^* = R_j - p - C_j \frac{n_j^* + 1}{\mu_2} \quad (31)$$

6. 结论

本文研究了一类具有两类平行顾客的经典双端排队模型, 该系统模拟乘客与出租车之间的供需匹配过程。通过构建一维马尔可夫链, 推导出系统稳态概率的解析表达式, 进而得到平均队列长度、平均等待时间等关键性能指标。社会收益函数综合了顾客净收益与出租车运营收益, 用于评估系统整体效率。数值分析表明, 顾客到达率、出租车服务率以及奖励成本结构均对等待时间和系统阈值产生显著影响, 例如提高空闲期出租车到达率可有效降低拥堵, 而奖励增加则促使顾客更倾向于加入系统。此外, 本文通过构建最优收费模型, 揭示了如何通过价格调控(如拥堵调节费或补贴)将顾客的个体决策引导至社会最优状态。这为平台运营商设计动态定价策略提供了理论依据。该研究为理解双端排队系统中顾客策略行为与社会优化提供了理论框架, 并为后续引入重试行为的扩展模型奠定了基础。

参考文献

- [1] Kendall, D.G. (1951) Some Problems in the Theory of Queues. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **13**, 151-173. <https://doi.org/10.1111/j.2517-6161.1951.tb00080.x>
- [2] Dobbie, J.M. (1961) Letter to the Editor—A Doubled-Ended Queueing Problem of Kendall. *Operations Research*, **9**, 755-757. <https://doi.org/10.1287/opre.9.5.755>
- [3] Kashyap, B.R.K. (1966) The Double-Ended Queue with Bulk Service and Limited Waiting Space. *Operations Research*, **14**, 822-834. <https://doi.org/10.1287/opre.14.5.822>
- [4] Manou, A., Economou, A. and Karaesmen, F. (2014) Strategic Customers in a Transportation Station: When Is It Optimal to Wait? *Operations Research*, **62**, 910-925. <https://doi.org/10.1287/opre.2014.1280>
- [5] Shi, Y. and Lian, Z. (2016) Optimization and Strategic Behavior in a Passenger-Taxi Service System. *European Journal of Operational Research*, **249**, 1024-1032. <https://doi.org/10.1016/j.ejor.2015.07.031>
- [6] Shi, Y. and Lian, Z. (2016) Equilibrium Strategies and Optimal Control for a Double-Ended Queue. *Asia-Pacific Journal of Operational Research*, **33**, Article 1650022. <https://doi.org/10.1142/s0217595916500226>
- [7] Bu, Q. and Sun, Y. (2024) Stationary Analyses for a Double-Ended Queueing System with Random Service Capacity and Balking Customers. *Journal of Computational and Applied Mathematics*, **437**, Article 115473. <https://doi.org/10.1016/j.cam.2023.115473>
- [8] Liu, X., Xu, X. and Liu, M. (2024) Strategy Analysis of Retrial Queue with Parallel Customer and Standby Server.

Methodology and Computing in Applied Probability, **26**, Article No. 27. <https://doi.org/10.1007/s11009-024-10098-7>

- [9] 张淞钛, 徐秀丽. 具有两类平行顾客的不完全故障排队系统均衡分析[J]. 系统科学与数学, 2019, 39(4): 637-647.
- [10] 徐秀丽, 张荣荣. 具有两类平行顾客的灾难清空排队系统均衡分析[J]. 河南师范大学学报(自然科学版), 2024, 52(1): 42-50.