

具有两类顾客和同步多重工作休假的可修排队

丁栋山, 李远鹏

燕山大学理学院统计系, 河北 秦皇岛

收稿日期: 2026年1月1日; 录用日期: 2026年1月28日; 发布日期: 2026年2月5日

摘要

本文研究了一个具有两类顾客、同步多重工作休假以及可修故障机制的可修排队系统。为更贴合现实服务环境, 模型假设两类顾客均服从泊松到达过程, 服务时间在正常与休假模式下均服从指数分布, 系统采用同步休假策略, 并考虑服务台在正常忙期发生的指数型故障与维修过程。基于系统状态转移率矩阵构建出三维连续时间马尔可夫链, 并证明其生成元矩阵具有拟生灭(QBD)结构。利用矩阵几何解法推导系统的稳态分布, 从而得到系统平均队长、平均逗留时间、休假概率、忙期长度及可用度等性能指标的解析表达式。通过数值实验分析模型参数对系统性能的敏感性, 并进一步构建个人效益函数与系统整体效益函数, 探讨最优参数设置。结果表明, 提高优先级顾客的服务率并适当增加服务台数量能够有效提升系统整体效益。

关键词

排队论, 非抢占优先权, 马尔科夫链, 矩阵几何解, 效益分析

A Repairable Queueing Model with Two Types of Customers and Synchronous Multiple Working Vacations

Dongshan Ding, Yuanpeng Li

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

Received: January 1, 2026; accepted: January 28, 2026; published: February 5, 2026

Abstract

Queueing models with two classes of customers of different priority levels are widely observed in real-life scenarios, for instance, priority users in ride-hailing services, emergency patients in hospitals and vehicles with prior reservations for charging stations. Consequently, many scholars at home

and abroad have incorporated dual customer classes as a key strategy in the construction and analysis of queueing models. In this paper, based on possible situations that may occur in real-world queueing systems, non-preemptive priority policy, multiple working vacation policy and repairable server policy are incorporated into the queueing model to construct a system that accurately reflects practical conditions. The main contents of the paper are as follows: model description of the state transition rate matrix of the system, analysis of the three-dimensional Markov chain, and the derivation of the system's steady-state distribution using the matrix-geometric method and iterative techniques. Based on the steady-state results, performance measures are derived and numerically analyzed. Finally, individual and social benefit functions are established for further analysis, and suggestions for improving overall efficiency are also proposed.

Keywords

Queueing Theory, Non-Preemptive Priority, Markov Chains, Matrix-Geometric Solution, Benefit Analysis

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

具有两类不同优先级顾客的排队模型在实际生活问题中普遍存在，例如网约车服务的优先用户、医院急诊病人、提前预约充电桩的车辆等，因此国内外有许多学者将两类顾客作为主要策略加入排队模型。

排队论起源于 20 世纪初丹麦工程师 A.K. Erlang [1] 对电话交换系统的开创性研究，其通过概率论方法解决了自动电话交换机的中继线计算问题。此后，随着 Kendall 符号的提出以及对非抢占优先权、服务器休假等复杂机制的深入分析，排队论逐渐演变为运筹学中研究随机服务系统性能的重要数学分支。在基础模型与机制研究方面，Kamoun [2] 分析了受相关马尔可夫中断过程影响的非抢占优先权排队系统的性能；朱翼隽[3]等探讨了多重工作休假的 $M/M/c$ 排队系统；Afnan [4] 等则进一步研究了带有服务器休假的非抢占优先权 $M/M/m$ 模型。针对复杂的顾客行为，Korenevskaya [5] 等讨论了带有随机挤出机制和非抢占优先权的重试排队系统；王慧[6]分析了带不耐烦顾客和单重休假的非抢占优先权模型；汪文杰[7]等则从博弈论视角分析了带有启动时间和故障的 $M/M/1$ 系统均衡策略。在特定模型演进与应用方面，王哲[8]将非抢占优先权排队理论应用于医疗服务系统的优化研究，郑晓铭[9]则在此基础上，进一步深入探讨了带抢占优先权和服务台休假的排队系统及其相关性能指标。针对具有两类优先权顾客的 $M/M/1$ 排队，张怡通[10]等给出了系统的优化分析。此外，关于带有接触匹配机制的 $M/M/c+m$ 模型，谭玉廷[11]等利用拟生灭过程分析了不同优先级顾客的匹配性能指标，为复杂服务系统的设计提供了理论支撑。

基于上述研究，本文研究了一个具有两类顾客、同步多重工作休假以及可修故障机制的排队系统。文章的第 2 节为模型描述，第 3 节为系统稳态分析以及矩阵分解过程，第 4 节为数值分析以及系统效益函数分析并给出了结论。

2. 模型描述

本文建立了一个具有两类顾客和同步多重工作休假且故障可修的 $M/M/c$ 排队模型，该模型将系统中的顾客分为 I 类和 II 类，其中 II 类顾客拥有更高的优先级别。模型假定两类顾客按照单队列排队，其中 I 类顾客可以无限到达，拥有无限的容量空间，II 类顾客最大容量空间为 c 。模型中排队的两类顾客均服从先

到先服务的排队规则。

具有两类顾客和同步多重工作休假的可修 $M/M/c$ 排队系统的具体模型描述如下:

(1) 到达过程: 两类顾客的到达均为 Possion 过程, 系统中 I 类顾客和 II 类顾客的到达时间间隔 T_1 和 T_2 分别服从参数为 λ_1 和 λ_2 的负指数分布, 其中 $\lambda_1, \lambda_2 > 0$ 。即 $P\{T_1 \leq x\} = 1 - e^{-\lambda_1 x}, x > 0$, $P\{T_2 \leq x\} = 1 - e^{-\lambda_2 x}, x > 0$ 。

(2) 服务过程: 系统在正规忙期为 I 类顾客和 II 类顾客的服务时间间隔 S_1 和 S_2 分别服从参数为 μ_1 和 μ_2 的负指数分布, 其中 $\mu_1, \mu_2 > 0$ 。即 $P\{S_1 \leq x\} = 1 - e^{-\mu_1 x}, x > 0$, $P\{S_2 \leq x\} = 1 - e^{-\mu_2 x}, x > 0$ 。

(3) 工作休假过程: 服务台采用首次休假为同步工作休假的多重休假模式, 若系统中服务台均为空时, 系统自动休假; 若在休假期间有顾客到达, 则系统休假时服务台以较低的服务速率 μ_w 为顾客服务, 其中服务台工作休假期间对 I 类顾客和 II 类顾客服务时间间隔 S_{w1} 和 S_{w2} 分别服从参数为 μ_{w1} 和 μ_{w2} 的负指数分布, 其中 $\mu_{w1}, \mu_{w2} > 0$ 。即 $P\{S_{w1} \leq x\} = 1 - e^{-\mu_{w1} x}, x > 0$, $P\{S_{w2} \leq x\} = 1 - e^{-\mu_{w2} x}, x > 0$ 。当一个假期结束时, 若有顾客在等待则服务台立即停止休假开始服务, 若系统内没有顾客等待则继续休假。系统工作休假结束后立即转为正常的速率 μ 为顾客服务。休假时间长度为 M 服从参数为 θ 的负指数分布。即 $P\{M \leq x\} = 1 - e^{-\theta x}, x > 0$ 。

(4) 故障维修过程: 系统为串联结构, 同步故障。在工作休假期间, 服务台仅在正规忙期会发生故障, 故障期间不进行服务, 所有服务台寿命 L 服从参数为 α 的负指数分布。系统中有 c 个修理工, 当服务台出现故障时立刻进行维修, 维修时间 U 服从参数为 β 的负指数分布, 维修后的机器恢复至全新状态, 即 $P\{L \leq x\} = 1 - e^{-\alpha x}, x > 0$, $P\{U \leq x\} = 1 - e^{-\beta x}, x > 0$ 。

(5) 排队规则: 在此排队系统中, 当 II 类顾客进入系统时, 若有系统有空闲服务台则直接接受服务; 若系统无空闲服务台且被 I 类顾客使用, 则 II 类顾客排队至 I 类顾客后等待服务。当系统中 II 类顾客超过 c 个, 则新到达的 II 类顾客消失。

模型中所涉及的变量均相互独立且服从负指数分布。

3. 系统稳态分析

3.1. 模型分析

令 $L_1(t)$ 和 $L_2(t)$ 表示为系统在 t 时刻时 I 类顾客和 II 类顾客的顾客数, $J(t)$ 为系统在 t 时刻的状态, 状态如下:

$$J(t) = \begin{cases} 0, & \text{时刻 } t \text{ 系统处于工作休假状态} \\ 1, & \text{时刻 } t \text{ 系统处于正规忙期状态} \\ 2, & \text{时刻 } t \text{ 系统处于故障状态} \end{cases}$$

则 $\{L_1(t), L_2(t), J(t), t \geq 0\}$ 是一个三维 Markov 链, 状态空间为:

$\Omega = \{(0, 0, 0) \cup \{(0, j, k), 1 \leq j \leq c, k = 0, 1, 2\} \cup \{(i, j, k), i \geq 1, 0 \leq j \leq c, k = 0, 1, 2\}\}$ 。其中状态空间 $(0, 0, 0)$ 表示系统处于工作休假状态且没有顾客; $(0, j, k)$ 状态下系统中只有 II 类顾客, 其中 $1 \leq j \leq c$; (i, j, k) 状态表示系统中既有 I 类顾客也有 II 类顾客, 其中 $i \geq 1$ 或 $0 \leq j \leq c$, $k = 0, 1, 2$ 分别表示系统处于工作休假、正规忙期和故障状态。将 $(0, 0, 0), (0, 1, 0), (0, 1, 1), (0, 1, 2), \dots, (0, c, 0), (0, c, 1), (0, c, 2)$ 定义为 0 水平, 将 $(i, 0, 0), (i, 0, 1), (i, 0, 2), (i, 1, 0), (i, 1, 1), (i, 1, 2), \dots, (i, c, 0), (i, c, 1), (i, c, 2)$ 定义为 i 水平。

3.2. 系统稳态条件分析

将系统状态空间按照字典序排列, 可以得到该 Markov 过程的状态概率转移矩阵 Q 如下:

$$Q = \begin{pmatrix} A_{00} & C_{01} & & & & & \\ B_{10} & A_{11} & C & & & & \\ & B_{21} & A_{22} & C & & & \\ & & B_{32} & A_{33} & C & & \\ & & & \ddots & \ddots & & \\ & & & & B_{c-1,c-2} & A_{c-1,c-1} & C \\ & & & & & B_{c,c-1} & A_{cc} & C \\ & & & & & & \ddots & \ddots \end{pmatrix} \quad (1)$$

其中 $A_{00}, A_{11}, \dots, A_{cc}; B_{10}, B_{21}, \dots, B_{c,c-1}; C_{01}, C$ 分别表示对应水平之间的状态转移率矩阵。为使矩阵更加简洁易读, 定义 $\xi_i, \xi_{iw}, \xi_c, \xi_{cw}, \xi_{i0}, \xi_{i0w}, \xi_{ij}, \xi_{ijw}, \xi_{ic}, \xi_{icw}, \phi, \phi_c$ 如下:

$$\begin{aligned} \xi_i &= -(\lambda_1 + \lambda_2 + i\mu_2 + i\alpha), 1 \leq i \leq c-1 \\ \xi_{iw} &= -(\lambda_1 + \lambda_2 + \theta + i\mu_{w2}), 1 \leq i \leq c-1 \\ \xi_c &= -(\lambda_1 + c\mu_2 + c\alpha) \\ \xi_{cw} &= -(\lambda_1 + \theta + c\mu_{w2}) \\ \xi_{i0} &= -(\lambda_1 + \lambda_2 + i\mu_1 + \alpha), 1 \leq i \leq c-1 \\ \xi_{i0w} &= -(\lambda_1 + \lambda_2 + \theta + i\mu_{w1}), 1 \leq i \leq c-1 \\ \xi_{ij} &= -(\lambda_1 + \lambda_2 + i\mu_1 + j\mu_2 + j\alpha), 1 \leq i \leq c-1, 1 \leq j \leq c-1 \\ \xi_{ijw} &= -(\lambda_1 + \lambda_2 + \theta + i\mu_{w1} + j\mu_{w2}), 1 \leq i \leq c-1, 1 \leq j \leq c-1 \\ \xi_{ic} &= -(\lambda_1 + i\mu_1 + c\mu_2 + c\alpha), 1 \leq i \leq c-1 \\ \xi_{icw} &= -(\lambda_1 + \theta + i\mu_{w1} + c\mu_{w2}), 1 \leq i \leq c-1 \\ \phi &= -(\lambda_1 + \lambda_2 + \beta) \\ \phi_c &= -(\lambda_1 + \beta) \end{aligned} \quad (2)$$

具体矩阵如下:

A_{00} 为 $(3c+1)$ 阶方阵, 即

$$A_{00} = \begin{pmatrix} -(\lambda_1 + \lambda_2) & \lambda_2 & & & & & & & \\ \mu_{w2} & \xi_{1w} & \theta & 0 & \lambda_2 & & & & \\ \mu_2 & 0 & \xi_1 & \alpha & 0 & \lambda_2 & & & \\ & & \beta & \phi & 0 & 0 & \lambda_2 & & \\ 2\mu_{w2} & 0 & 0 & \xi_{2w} & \theta & 0 & \lambda_2 & & \\ 2\mu_2 & 0 & 0 & \xi_2 & 2\alpha & 0 & \lambda_2 & & \\ & & & \beta & \phi & 0 & 0 & \lambda_2 & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & (c-1)\mu_{w2} & 0 & 0 & \xi_{(c-1)w} & \theta & 0 & \lambda_2 \\ & & & (c-1)\mu_2 & 0 & 0 & \xi_{c-1} & (c-1)\alpha & 0 & \lambda_2 \\ & & & & & & \beta & \phi & 0 & 0 & \lambda_2 \\ & & & & & & c\mu_{w2} & 0 & 0 & \xi_{cw} & \theta & 0 \\ & & & & & & c\mu_2 & 0 & 0 & \xi_c & c\alpha & \\ & & & & & & & & & \beta & \phi_c \end{pmatrix} \quad (3)$$

A_{ii} 为 $(3c+3)$ 阶方阵, 即

$$A_{ii} = \left(\begin{array}{cccccc} \xi_{i0w} & \theta & 0 & \lambda_2 & & \\ \xi_{i0} & \alpha & 0 & \lambda_2 & & \\ \beta & 0 & \phi & 0 & \lambda_2 & \\ \mu_{w2} & 0 & 0 & \xi_{i1w} & \theta & \lambda_2 \\ \mu_2 & 0 & 0 & \xi_{i1} & \alpha & 0 & \lambda_2 \\ & & & \beta & \phi & 0 & 0 & \lambda_2 \\ & 2\mu_{w2} & 0 & 0 & \xi_{i2w} & \theta & 0 & \lambda_2 \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & (c-1)\mu_{w2} & 0 & 0 & \xi_{i(c-1)w} & \theta & 0 & \lambda_2 \\ & (c-1)\mu_2 & 0 & 0 & \xi_{i(c-1)} & (c-1)\alpha & 0 & \lambda_2 \\ & & & & \beta & \phi & 0 & 0 & \lambda_2 \\ & c\mu_{w2} & 0 & 0 & \xi_{icw} & \theta & 0 \\ & c\mu_2 & 0 & 0 & \xi_{ic} & c\alpha \\ & & & & \beta & \phi_c \end{array} \right), 1 \leq i \leq c \quad (4)$$

B_{10} 为 $(3c+3) \times (3c+1)$ 阶方阵, 即

$$B_{10} = \left(\begin{array}{cccccc} \mu_{w1} & & & & & & \\ \mu_1 & & & & & & \\ 0 & & & & & & \\ \mu_{w1} & & & & & & \\ & \mu_1 & & & & & \\ & & 0 & & & & \\ & & & \mu_{w1} & & & \\ & & & & \ddots & & \\ & & & & & 0 & \\ & & & & & & \mu_{w1} \\ & & & & & & & \mu_1 \\ & & & & & & & & 0 \end{array} \right) \quad (5)$$

$B_{i,i-1}$ 为 $(3c+3) \times (3c+3)$ 阶方阵, 即

$$B_{i,i-1} = \left(\begin{array}{cccccc} i\mu_{w1} & & & & & & \\ i\mu_1 & & & & & & \\ 0 & & & & & & \\ i\mu_{w1} & & & & & & \\ & i\mu_1 & & & & & \\ & & 0 & & & & \\ & & & i\mu_{w1} & & & \\ & & & & i\mu_1 & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & i\mu_{w1} \\ & & & & & & & & i\mu_1 \\ & & & & & & & & & 0 \end{array} \right), 2 \leq i \leq c \quad (6)$$

C_{01} 为 $(3c+1) \times (3c+3)$ 阶方阵, 即

$$\mathbf{C}_{01} = \begin{pmatrix} \lambda_1 & & & & & \\ 0 & 0 & \lambda_1 & & & \\ & & & \lambda_1 & & \\ & & & & \ddots & \\ & & & & & \lambda_1 \\ & & & & & & \lambda_1 \end{pmatrix} \quad (7)$$

\mathbf{C} 为 $(3c+3) \times (3c+3)$ 阶方阵, 即

$$\mathbf{C} = \begin{pmatrix} \lambda_1 & & & & & \\ & \lambda_1 & & & & \\ & & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \lambda_1 \end{pmatrix} \quad (8)$$

通过分析 \mathbf{Q} 矩阵的结构特征可知, 此三维 Markov 状态过程 $\{L_1(t), L_2(t), J(t), t \geq 0\}$ 是拟生灭过程 (Quasi-Birth-and-Death process, QBD)。

令矩阵 $\mathbf{H} = \mathbf{A}_{cc} + \mathbf{B}_{c,c-1} + \mathbf{C}$ 得

$$\mathbf{H} = \begin{pmatrix} -(\lambda_2 + \theta) & \theta & 0 & \lambda_2 & & & & \\ & -(\lambda_2 + \alpha) & \alpha & 0 & \lambda_2 & & & \\ \beta & 0 & -(\lambda_2 + \beta) & 0 & 0 & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ c\mu_{w2} & 0 & 0 & -(\theta + c\mu_{w2}) & \theta & 0 & & \\ c\mu_2 & 0 & 0 & 0 & -(\alpha + c\mu_2) & c\alpha & & \\ & & & & \beta & -\beta & & \end{pmatrix} \quad (9)$$

令 $\pi = \{\pi_1, \pi_2, \pi_3, \dots, \pi_c\}$ 表示为 H 的稳态概率向量, 则满足以下方程组

$$\begin{cases} \pi H = 0 \\ \pi e = 1 \end{cases} \quad (10)$$

由上述平衡方程可得下列方程组

$$-(\lambda_2 + \theta)\pi_0 + \beta\pi_2 + \mu_{w2}\pi_3 = 0 \quad (11)$$

$$\theta\pi_0 - (\lambda_2 + \alpha)\pi_1 + \mu_2\pi_4 = 0 \quad (12)$$

$$\alpha\pi_1 - (\lambda_2 + \beta)\pi_2 = 0 \quad (13)$$

$$-(\lambda_2 + \theta + i\mu_{w2})\pi_{i+2} + \lambda_2\pi_{i-1} + (i+1)\mu_{w2}\pi_{i+5} = 0, \quad 1 \leq i \leq c-1 \quad (14)$$

$$-(\lambda_2 + j\alpha + j\mu_2)\pi_{j+3} + \lambda_2\pi_j + j\mu_2\pi_{j+5} = 0, \quad 1 \leq j \leq c-1 \quad (15)$$

$$-(\lambda_2 + \beta)\pi_{k+4} + \lambda_2\pi_{k+1} + k\alpha\pi_{k+3} = 0, \quad 1 \leq k \leq c-1 \quad (16)$$

$$-(\theta + c\mu_{w2})\pi_{c-2} + \lambda_2\pi_{c-5} = 0 \quad (17)$$

$$-(c + c\mu_2)\pi_{c-1} + \lambda_2\pi_{c-4} + \theta\pi_{c-2} + \beta\pi_c = 0 \quad (18)$$

$$-c\alpha\pi_{c-1} + \lambda_2\pi_{c-3} - \beta\pi_c = 0 \quad (19)$$

$$\pi_0 e + \sum_{i=1}^c \pi_i e = 1 \quad (20)$$

通过上述式子递推计算可以求解并得到各概率 $\{\pi_1, \pi_2, \pi_3, \dots, \pi_c\}$ 。

根据拟生灭过程 $\{L_1(t), L_2(t), J(t), t \geq 0\}$ 正常返的充分必要条件 $\pi C e < \pi B_{c,c-1} e$ 可得

$$\begin{pmatrix} c\mu_{w1} & & & & \\ & c\mu_1 & & & \\ & & 0 & & \\ & & & \ddots & \\ (\pi_1, \pi_2, \dots, \pi_c) & & & & \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} > (\pi_1, \pi_2, \dots, \pi_c) \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_1 & & & \\ & & \ddots & & \\ & & & \lambda_1 & \\ & & & & \lambda_1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \quad (21)$$

故系统稳态平稳条件可以表示为

$$\lambda_1 < \frac{c(\mu_{w1} + \mu_1)}{3} \quad (22)$$

当满足上述条件时, 该 QBD 过程是正常返的。

稳态概率的矩阵几何解法

当三维 Markov 过程 $\{L_1(t), L_2(t), J(t), t \geq 0\}$ 为正常返时, 稳态分布定义如下:

$$P_{i,j,k} = \lim_{t \rightarrow \infty} P\{L_1(t) = i, L_2(t) = j, J(t) = k\}, (i, j, k) \in \Omega$$

$$P = (P_0, P_1, P_2, \dots, P_c, P_{c+1}, \dots)$$

$$P_0 = (P_{0,0,0}, P_{0,1,0}, P_{0,1,1}, P_{0,1,2}, \dots, P_{0,c,0}, P_{0,c,1}, P_{0,c,2})$$

$$P_i = (P_{i,0,0}, P_{i,0,1}, P_{i,0,2}, P_{i,1,0}, P_{i,1,1}, P_{i,1,2}, \dots, P_{i,c,0}, P_{i,c,1}, P_{i,c,2}), i \geq 1$$

稳态概率向量 P 满足方程组

$$\begin{cases} PQ = 0 \\ Pe = 1 \end{cases} \quad (23)$$

将方程组展开可得

$$P_0 A_{00} + P_1 B_{10} = 0 \quad (24)$$

$$P_0 C_{01} + P_1 A_{11} + P_2 B_{21} = 0 \quad (25)$$

$$P_i C + P_{i+1} A_{i+1,i+1} + P_{i+2} B_{i+2,i+1} = 0, 1 \leq i \leq c-2 \quad (26)$$

$$P_{c-1} C + P_c A_{cc} = 0 \quad (27)$$

$$P_c C = 0 \quad (28)$$

$$P_0 e + \sum_{i=1}^c P_i e = 1 \quad (29)$$

其中 e 为适当维数且元素全为 1 的列向量。通过对方程组联立求解可以得到稳态概率向量 $P = (P_0, P_1, P_2, \dots, P_c)$ 。

三维 Markov 过程 $\{L_1(t), L_2(t), J(t), t \geq 0\}$ 正常返的充要条件为矩阵二次方程 $R^2 B_{c,c-1} + RA_{cc} + C = 0$

存在最小非负解 R 且其最小谱半径 $sp(R) < 1$, 同时构建的 $3c^2 + 6c + 1$ 维方阵

$$B[R] = \begin{pmatrix} A_{00} & C_{01} & & & & \\ B_{10} & A_{11} & C & & & \\ B_{21} & A_{22} & C & & & \\ & B_{32} & A_{33} & C & & \\ & & \ddots & \ddots & \ddots & \\ & & & B_{c-1,c-2} & A_{c-1,c-1} & C \\ & & & B_{c,c-1} & RA_{c,c-1} + A_{cc} & \end{pmatrix} \quad (30)$$

存在左零向量。当该过程正常返时, 系统稳态分布满足下列方程组:

$$\begin{cases} (P_0, P_1, P_2, \dots, P_c) B[R] = 0 \\ \sum_{i=0}^{c-1} P_i e + P_c (I - R)^{-1} e = 1 \\ P_i = P_c R^{i-c}, i \geq c \end{cases} \quad (31)$$

证明: 根据 Q 矩阵以及 $B[R]$ 矩阵结构可知, $B[R]$ 矩阵是一个 $3c^2 + 6c + 1$ 维的方阵。首先, 由 $(P_0, P_1, P_2, \dots, P_c) B[R] = 0$ 可知 $B[R]$ 存在左零向量; 其次, 该三维 Markov 过程满足平衡方程 $PQ = 0$, 带入化简即可得 $P_i = P_c R^{i-c}, i \geq c$ 。

最后通过正规化条件 $P_0 e + \sum_{i=1}^{\infty} P_i e = 1$ 以及 $P_i = P_c R^{i-c}, i \geq c$ 可得

$$P_0 e + \sum_{i=1}^{\infty} P_i e = P_0 e + P_1 e + \dots + P_c e + P_{c+1} e + \dots = \sum_{i=0}^{c-1} P_i e + P_c \sum_{k=0}^{\infty} R^k e \quad (32)$$

由于 R 矩阵的最小谱半径 $sp(R) < 1$, 所以级数 $\sum_{k=0}^{\infty} R^k$ 收敛于 $(I - R)^{-1}$, 带入上式可推导出 $\sum_{i=0}^{c-1} P_i e + P_c (I - R)^{-1} e = 1$ 。其中 e 为适当维数的列向量且元素全为 1, I 为 $3c + 3$ 维单位阵。

4. 系统性能分析

4.1. 系统稳态性能指标

通过上述分析可以得到系统的稳态概率, 根据系统的稳态概率可求出系统的各项性能指标, 具体性能指标表达式如下:

(1) I类顾客平均队长为

$$E(L_1) = \sum_{i=1}^{\infty} i P(L_1 = i) = \sum_{i=1}^{\infty} \sum_{j=0}^c \sum_{k=0}^1 i P_{i,j,k} \quad (33)$$

(2) II类顾客平均队长为

$$E(L_2) = \sum_{j=1}^{\infty} j P(L_2 = j) = \sum_{j=1}^c \sum_{i=0}^{\infty} \sum_{k=0}^1 j P_{i,j,k} \quad (34)$$

(3) 系统内顾客平均逗留时间为

$$E(W) = \frac{1}{\lambda_1 + \lambda_2} [E(L_1) + E(L_2)] = \frac{1}{\lambda_1 + \lambda_2} \left(\sum_{i=1}^{\infty} \sum_{j=0}^c \sum_{k=0}^1 i P_{i,j,k} + \sum_{j=1}^c \sum_{i=0}^{\infty} \sum_{k=0}^1 j P_{i,j,k} \right) \quad (35)$$

(4) 服务台工作休假的概率为

$$P_v = \sum_{i=0}^{\infty} \sum_{j=0}^c p_{i,j,0} \quad (36)$$

(5) 系统平均忙期为

$$E(B) = \frac{1}{\mu_1} E(L_1) + \frac{1}{\mu_2} E(L_2) = \frac{\mu_2 E(L_1) + \mu_1 E(L_2)}{\mu_1 \mu_2} \quad (37)$$

(6) 服务台可用度为

$$P_{use} = \frac{\min\{E(L_1) + E(L_2), c\}}{c} \quad (38)$$

4.2. 数值分析

通过调整系统输入参数得到各性能指标随参数变化的结果并将结果进行可视化处理, 分析参数对性能指标变化趋势的影响。

图 1 中固定参数 $\lambda_1 = 1.8, \lambda_2 = 2.5, \mu_1 = 2, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1$ 研究系统在正规忙期对 II 类顾客的服务率 μ_2 以及系统服务台数 c 对 I 类顾客平均队长 $E(L_1)$ 的影响。当 μ_2 固定不变时, I 类顾客的平均队长 $E(L_1)$ 随着服务台数量 c 的增加而增大。因为 II 类顾客具有更高的优先级, 系统中新增加的服务台会优先给予 II 类顾客使用, 服务台数量 c 的增加会使系统中 II 类顾客接受服务的机会增大, 此时新到达的 I 类顾客就会在系统中等待服务从而导致 I 类顾客的平均队长增加; 当 c 固定不变时, 系统中 I 类顾客的平均队长 $E(L_1)$ 随着 μ_2 的增大而减小。由于系统对 II 类顾客的服务率增加, 会使系统整体的服务率加快, 从而使服务台的翻台率更高, 在同一时间内一个服务台可以服务更多的 II 类顾客, 有更多的空闲服务台可以为 I 类顾客进行服务, 所以系统中 I 类顾客的平均队长 $E(L_1)$ 会随着 μ_2 的增大而减小。故在其他参数不变的情况下, 提高系统对 II 类顾客的服务率可以提高系统的运行效率, 减少系统内 I 类等待顾客的平均队长。在系统对 II 类顾客服务率不变的情况下, 盲目增加系统服务台的数量可能会导致系统内等待顾客增加从而加重系统负担。

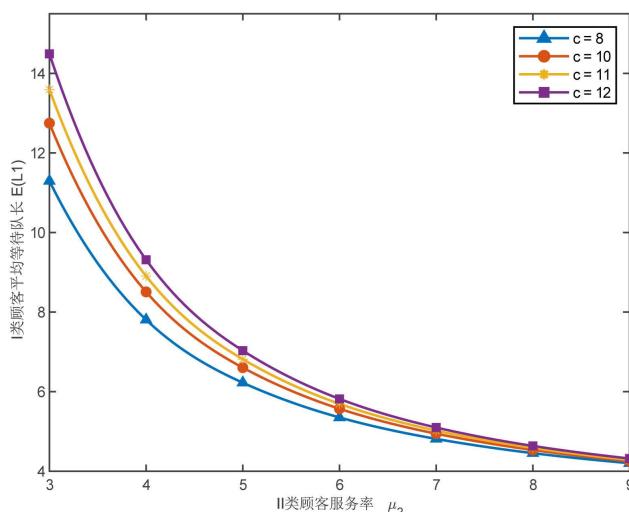


Figure 1. Relationship between the mean queue length $E(L_1)$, μ_2 , and c

图 1. 平均队长 $E(L_1)$ 和 μ_2 和 c 的关系

图 2 中固定参数 $\lambda_2 = 2.5, \mu_1 = 2, \mu_2 = 6, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, c = 10$ 展示了在不同维修速率 β 下 I 类顾客的平均队长 $E(L_1)$ 与服务台故障率 α 以及 I 类顾客到达率 λ_1 的变化关系。在其它参数均固定不变的情况下, I 类顾客的平均队长 $E(L_1)$ 随着维修速率的增大而减小。由于对故障系统的维修速度加快, 系统可以在更短的时间内重新为顾客提供服务, 因此提高对系统的维修速率可以使系统中 I 类顾客的平均队长减少; 当固定维修速率 β 不变时, I 类顾客的平均队长 $E(L_1)$ 随着服务台故障率 α 以及到达率 λ_1 的增大而增大。串联系统中一台服务台发生故障会导致系统整体停止运转, 服务台故障率越大, 服务台间歇停止服务的次数就越多, 间歇期间新到达的顾客就需要排队等待服务, 同时随着 I 类顾客到达率的 λ_1 的增大, 系统中 I 类顾客增加速率变快, 最后导致系统中 I 类顾客的平均队长增大。

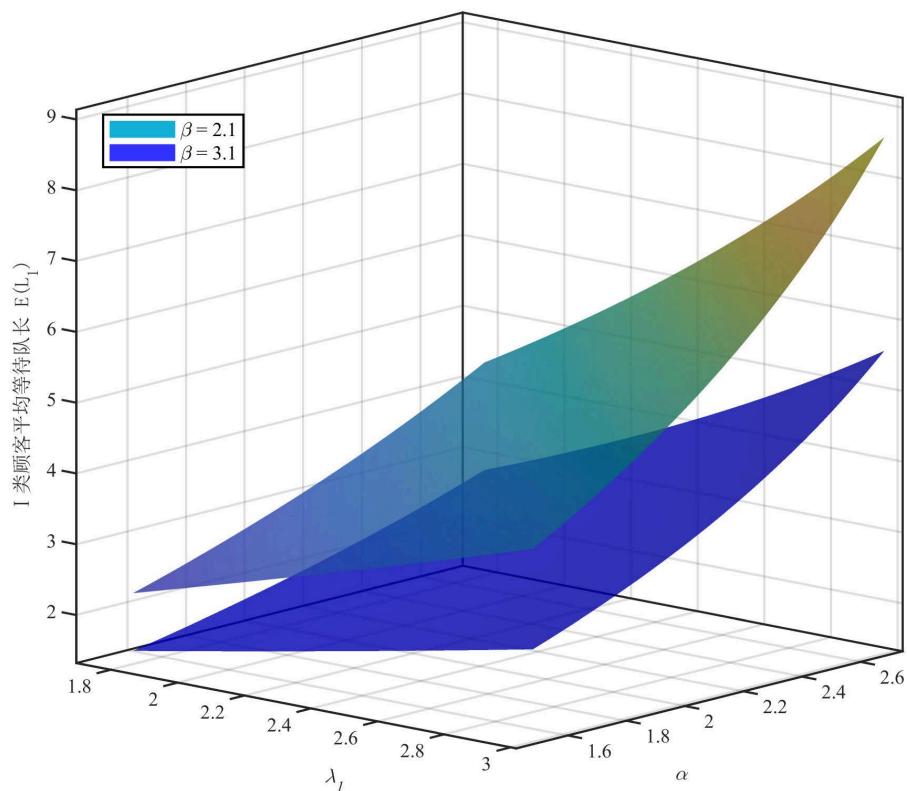


Figure 2. Relationship between the mean queue length $E(L_1)$, λ_1 , α , and β

图 2. 平均队长 $E(L_1)$ 和 λ_1 , α 和 β 的关系

图 3 中固定参数 $\lambda_1 = 1.8, \mu_1 = 2, \mu_2 = 6, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1, c = 10$ 分别展示了系统在设置服务台数量 $c = 10, c = 11, c = 12$ 情况下 II 类顾客平均队长 $E(L_2)$ 与 II 类顾客到达率 λ_2 的关系。当到达率 λ_2 固定不变时, II 类顾客平均队长 $E(L_2)$ 随着服务台数量 c 的增加而增大。这是因为在参数设定中 II 类顾客到达率大于 I 类顾客的到达率 ($\lambda_2 > \lambda_1$), 导致 II 类顾客形成自身拥堵现象, 同时服务台的数量增加使得 I 类顾客在多个服务台之间快速流动, II 类顾客打断 I 类顾客的机会变少, 非抢占优先级机制被削弱导致 II 类顾客等待服务台空闲的平均时间增加, 最终导致 II 类顾客的平均队长增加; 当服务台数量 c 固定不变时, II 类顾客平均队长 $E(L_2)$ 随着到达率 λ_2 的增大而增大。因为系统中服务台数量不变以及服务台对 II 类顾客服务速率不变时, 随着到达率 λ_2 的增大, 系统中的 II 类顾客会越来越多, 从而使 II 类顾客的平均队长 $E(L_2)$ 增加。

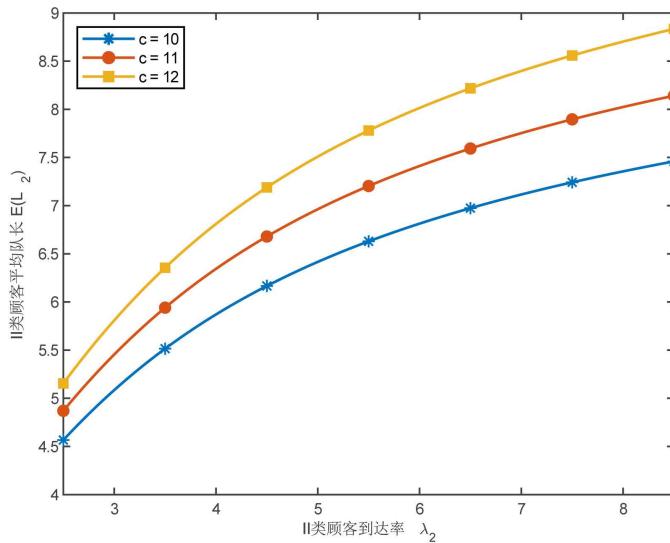
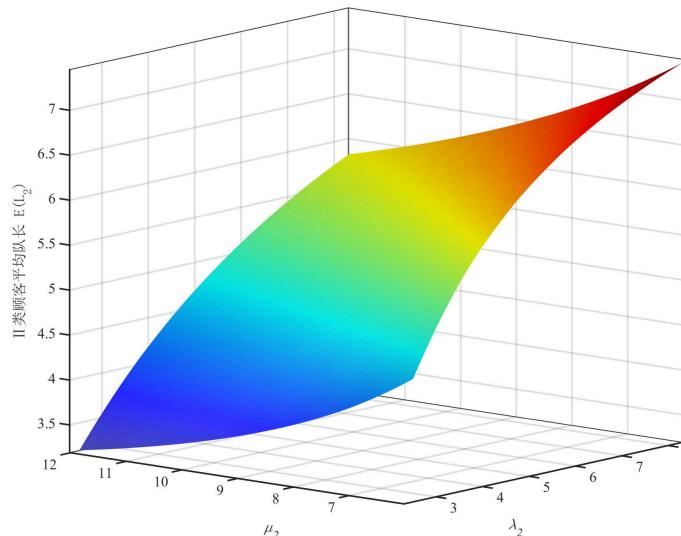
**Figure 3.** Relationship between the mean queue length $E(L_2)$, λ_2 , and c **图 3.** 平均队长 $E(L_2)$ 和 λ_2 和 c 的关系

图 4 中固定参数 $\lambda_1 = 1.8, \mu_1 = 2, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1, c = 10$ 研究系统中II类顾客到达率 λ_2 以及服务率 μ_2 对II类顾客平均队长 $E(L_2)$ 的影响。当 λ_2 固定不变时, 可以看到, II类顾客的平均队长 $E(L_2)$ 随着服务率 μ_2 的增大而减小。由于系统中其它参数不变, 提高服务台对II类顾客的服务率可以使正在排队的II类顾客减少排队等待时间, 更快的前往服务台接受服务, 从而使系统内II类顾客的数量减少, 缩短II类顾客的平均等待队长 $E(L_2)$; 当服务率 μ_2 固定不变时, II类顾客的平均队长 $E(L_2)$ 随着到达率 λ_2 的增大而增大。因为在服务台对II类顾客服务率不变的情况下, 到达率 λ_2 越大, 系统中进入的II类顾客就越多, 但是系统的负载能力不变, 新到达的II类顾客就需要排队等候服务, 最终导致II顾客的平均队长 $E(L_2)$ 随着到达率 λ_2 的增大而增大。

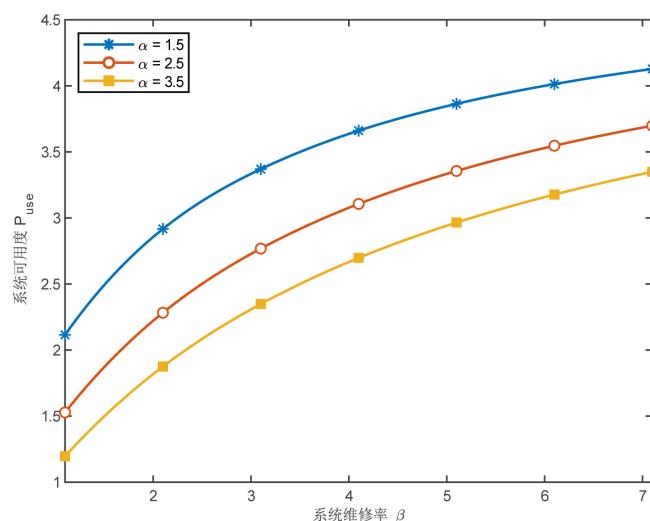
**Figure 4.** Relationship between the mean queue length $E(L_2)$, μ_2 , and λ_2 **图 4.** 平均队长 $E(L_2)$ 和 μ_2 和 λ_2 的关系

系统在固定参数 $\lambda_2 = 3.5, \mu_1 = 3.5, \mu_2 = 5, \mu_{w1} = 2, \mu_{w2} = 3, \theta = 5, \alpha = 1.5, \beta = 1.1$, 服务台数分别为 $c = 8, c = 9, c = 10$ 的情况下, 服务台工作休假概率 P_v 与 I 类顾客到达率 λ_1 关系如表 1 所示。从表 1 中可以得出如下结论: 系统服务台数量不变的情况下, 服务台工作休假的概率随着 I 类顾客到达率的增大而减小。主要原因是服务台为空时自动进行休假, 休假期间有顾客到达服务台则进入工作休假状态。系统中 I 类顾客到达率的增加会使得系统内有更多的顾客排队等待服务, 这会使空服务台占所有服务台的比例减少, 从而导致服务台进入休假状态的机会减少, 持续保持在正规忙期的状态; 当固定 I 类顾客到达率 λ_1 不变的情况下, 系统设置更多的服务台会使服务台工作休假概率 P_v 增加。由于系统在其他参数固定不变的情况下增加服务台的数量会提高系统整体的运行效率, 使得服务台空闲的概率增加, 空闲服务台自动进入休假状态, 提高了服务台工作休假概率 P_v 。

Table 1. Relationship between the working vacation probability P_v , c , and λ_1 **表 1.** 工作休假概率 P_v 随参数 c 和 λ_1 的关系

c	服务台工作休假概率 P_v				
	$\lambda_1 = 1.8$	$\lambda_1 = 2.8$	$\lambda_1 = 3.8$	$\lambda_1 = 4.8$	$\lambda_1 = 5.8$
8	0.5707	0.3283	0.182	0.0916	0.0356
9	0.597	0.3463	0.1961	0.1038	0.0468
10	0.6199	0.3616	0.2077	0.1135	0.0555

图 5 中描绘了系统在固定参数 $\lambda_1 = 1.8, \lambda_2 = 2.5, \mu_1 = 2, \mu_2 = 6, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, c = 10$ 的情况下, 系统故障率 α 以及维修率 β 对系统服务台可用度 P_{use} 的影响。可以发现, 当维修速率 β 不变时, 系统可用度 P_{use} 随着故障率的增大而减小; 在同一故障率水平下, P_{use} 随着 β 的增大而增大。这是因为系统的故障率影响服务台的工作时间, 当系统发生故障时, 系统不进行服务, 系统发生故障的概率越大, 服务停止工作的时间占比就越长, 导致系统可用度 P_{use} 减小; 反之, 若故障率确定的情况下, 对系统进行维修的速度越快, 系统恢复正常运转的速度就越快, 系统正常服务的时间占比就会增加, 使系统可用度增加。

**Figure 5.** Relationship between the server availability P_{use} , β , and α **图 5.** 服务台可用度 P_{use} 和 β , α 的关系

4.3. 效益函数

根据数值分析的结果可以发现, 系统性能指标对不同参数的变化有不同的敏感度, 为了分析系统的经济成本, 提高系统整体的运行效率, 本节通过建立个人效益函数以及系统效益函数探讨使系统收益达到最优的参数。

4.3.1. 个人效益函数

假定系统完成一次服务时, I类顾客和II类顾客的个人收益分别为 Z_1, Z_2 , 在系统中等待服务产生的损失为 Q_1, Q_2 。则两类顾客在系统中所获得的个人效益函数可以分别表示为

$$\begin{cases} W_1 = \mu_1 Z_1 - Q_1 \frac{E(L_1)}{\lambda_1} \\ W_2 = \mu_2 Z_2 - Q_2 \frac{E(L_2)}{\lambda_2} \end{cases} \quad (39)$$

图6中系统在固定参数 $Z_1 = 19, Q_1 = 3, \lambda_2 = 2.5, \mu_1 = 2, \mu_2 = 6, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1$ 的情况下, I类顾客个人效益 W_1 随到达率 λ_1 以及服务台数 c 的变化关系如图6所示。当服务台数 c 固定不变时, W_1 随着I类顾客到达率 λ_1 的增大而减少。由于I类顾客到达率的增加会使得系统内I类顾客的数量增加, 导致服务资源拥堵, 增加I类顾客的等待时间, 增加时间成本导致个人效益降低; 当固定I类顾客到达率 λ_1 不变时, 可以发现当 $\lambda_1 < 2.7$ 时, 增加服务台数会导致I类顾客的个人效益 W_1 减小, 这是因为新增加的服务台会优先服务系统中具有更高优先级别的II类顾客。当 $\lambda_1 \geq 2.7$ 时, 增加服务台数会导致I类顾客的个人效益 W_1 增大, 由于此时新增服务台会导致II类顾客形成自身拥堵现象, 同时服务台的数量增加使得I类顾客在多个服务台之间快速流动, 可以减少I类顾客等待时间, 所以新增服务台会使I类顾客的个人效益增加。

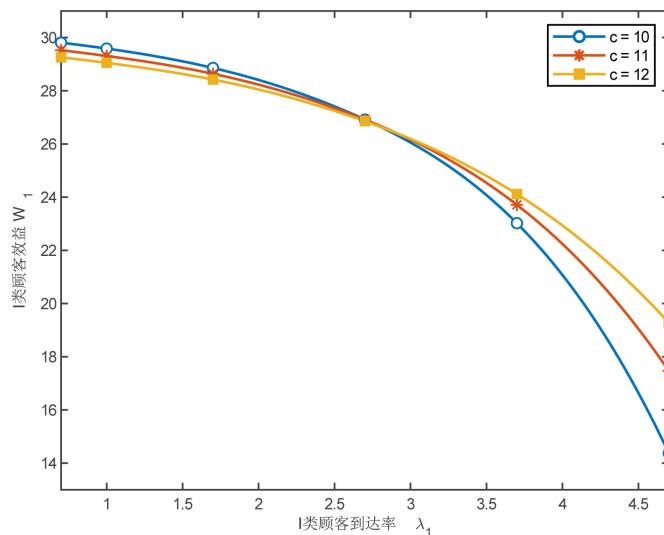


Figure 6. Relationship between the benefit W_1 , λ_1 , and c

图6. 效益 W_1 和 λ_1 和 c 的关系

表2中固定参数 $Z_2 = 15, Q_2 = 4, \lambda_2 = 1.8, \lambda_1 = 2.5, \mu_1 = 2, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1$, 研究系统分别在服务台数 $c = 10, c = 11, c = 12$ 情况下II类顾客个人效益 W_2 随服务率 μ_2 的变化情况。从表中数据可以发现当固定服务台数 c 时, W_2 随着 μ_2 的增大而增大。由于系统提高对II类顾客的服务率, 使单位时间内

有更多的II类顾客完成服务, 显著减少其等待时间, 且服务率增大使得II类顾客在系统中获得的收益增加。所以对于II类顾客来讲, 提高服务率 μ_2 可以显著提高其个人效益; 当固定服务率 μ_2 时, W_2 随着服务台数 c 的增加而小幅减少, 说明II类顾客个人效益的变化对服务台数 c 的变化有着较低的敏感度, 但是持续的增加服务台数量会使II类顾客形成自身拥堵现象, 导致其个人效益降低。

Table 2. Relationship between the individual benefit function W_2 , c , and μ_2 **表 2.** 个人效益函数 W_2 随参数 c 和 μ_2 的关系

c	个人效益函数 W_2						
	$\mu_2 = 4$	$\mu_2 = 5$	$\mu_2 = 6$	$\mu_2 = 7$	$\mu_2 = 8$	$\mu_2 = 9$	$\mu_2 = 10$
10	50.7084	66.8582	82.6938	98.3122	113.7826	129.1499	144.4437
11	49.9305	66.26	82.2084	97.901	113.4223	128.8263	144.1475
12	49.171	65.6884	81.7534	97.5208	113.0924	128.532	143.8792

4.3.2. 系统效益函数

从系统整体效益来看, 假定系统每服务完一名顾客得到的平均收益为 Z_p , 花费平均成本为 Q_p 。则系统的整体效益函数可以表示为

$$W_p = \lambda_p (\mu_p Z_p - Q_p E(D_p)) \quad (40)$$

其中

$$\lambda_p = \frac{\lambda_1 + \lambda_2}{2} \quad (41)$$

$$\mu_p = \frac{\mu_1 + \mu_2}{2} \quad (42)$$

$$E(D_p) = \frac{\lambda_2 E(L_1) + \lambda_1 E(L_2)}{2\lambda_1 \lambda_2} \quad (43)$$

系统在固定参数 $Z_p = 20, Q_p = 3.5, \lambda_1 = 1.8, \lambda_2 = 2.5, \mu_1 = 2, \mu_{w1} = 1, \mu_{w2} = 2, \theta = 5, \alpha = 1.5, \beta = 1.1$, 服务台数分别为 $c = 10, c = 11, c = 12$ 的情况下, 系统整体效益 W_p 与II类顾客服务率 μ_2 关系如表3所示。观察数据可以发现, 在系统服务台数量不变的情况下, 系统整体效益 W_p 随着 μ_2 的增大而增加。这是因为系统提高对II类顾客的服务率可以有效减少系统内顾客的等待时间, 使系统在单位时间内流通的顾客数量增加, 从而提高系统的整体效益; 当II类顾客服务率 μ_2 不变时, W_p 随着服务台数量的增加而减少。由于系统增加服务台数量会提高系统运行的成本, 从而使系统整体的收益降低。

Table 3. Relationship between the social benefit function W_p , parameter c , and μ_2 **表 3.** 系统整体效益函数 W_p 随参数 c 和 μ_2 的关系

c	系统整体效益 W_p						
	$\mu_2 = 4$	$\mu_2 = 5$	$\mu_2 = 6$	$\mu_2 = 7$	$\mu_2 = 8$	$\mu_2 = 9$	$\mu_2 = 10$
10	102.4734	129.0391	153.4856	176.8757	199.6746	222.1149	244.3235
11	100.9156	128.0325	152.7698	176.327	199.2298	221.7382	243.9945
12	99.3419	127.0446	152.0827	175.8085	198.8132	221.3882	243.69

5. 结论

综上所述, 可以通过增加单个服务台对II类顾客服务率 μ_2 , 并尽可能不增设系统内服务台的数量来提高系统整体效益。同时在保证系统整体效益的前提下控制I类顾客的到达率来保证I类顾客的个人效益。

参考文献

- [1] Erlang, A.K. (1909) The Theory of Probabilities and Telephone Conversations. *Electrotechnische Zeitschrift*, **30**, 482-485.
- [2] Kamoun, F. (2008) Performance Analysis of a Non-Preemptive Priority Queueing System Subjected to a Correlated Markovian Interruption Process. *Computers & Operations Research*, **35**, 3969-3988.
<https://doi.org/10.1016/j.cor.2007.06.001>
- [3] 朱翼隽, 徐剑, 周宗好. 多重工作休假的 $M/M/c$ 排队系统[J]. 江苏大学学报(自然科学版), 2012, 33(3): 369-372.
- [4] Afnan, M.I., Jain, M. and Sharma, R. (2022) Non-Preemptive Priority $M/M/m$ Queue with Servers' Vacations. *International Journal of Mathematics in Operations Research*, **21**, 120-137.
- [5] Korenevskaya, M., Zayats, O., Ilyashenko, A. and Muliukha, V. (2019) Retrial Queueing System with Randomized Push-Out Mechanism and Non-Preemptive Priority. *Procedia Computer Science*, **150**, 716-725.
<https://doi.org/10.1016/j.procs.2019.02.016>
- [6] 王慧. 带不耐烦顾客和单重休假的非抢占优先权排队模型分析[J]. 现代商贸工业, 2019, 40(32): 190-191.
- [7] 汪文杰, 刘力维, 叶晴晴. 带有启动时间和两类故障的 $M/M/1$ 系统均衡策略分析[J]. 重庆工商大学学报(自然科学版), 2022, 39(2): 38-45.
- [8] 王哲. 非抢占优先权排队在医疗服务系统中的应用研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2018.
- [9] 郑晓铭. 带抢占优先权和服务台休假的排队系统研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2017.
- [10] 张怡通, 徐秀丽. 具有两类优先权顾客的 $M/M/1$ 排队的优化分析[J]. 应用概率统计, 2021, 37(5): 449-460.
- [11] 谭玉廷, 徐秀丽, 郑瑞. 带有非抢占优先权接触匹配的 $M/M/c+m$ 排队模型分析[J]. 山东大学学报(理学版), 2023, 58(11): 45-52.