

# 考虑顾客耐心和实时座席转移的联络中心运营指标优化研究

李成龙, 李军祥, 周婷婷

(上海理工大学 管理学院, 上海 200093)

**摘要:** 为解决传统联络中心服务不同类型顾客产生的座席不平衡问题, 提高联络中心的服务质量和优化顾客服务体验, 通过引入部分多技能座席, 在不增加座席的情况下根据队列长度与座席状态对部分多技能座席进行实时转移, 考虑顾客放弃的条件下以顾客的放弃率最低为目标构建出一个新的联络中心模型。运用 Arena 仿真软件建立模型对新联络中心模型的整体性能进行仿真分析, 并与传统模型进行对比验证其有效性。结果表明, 新的联络中心模型在提高了接通率的同时减少了顾客的排队等待时间, 平衡了座席的服务强度。

**关键词:** 联络中心; 多服务; 多技能; 实时转移; 耐心值

中图分类号: F 273.1 文献标志码: A

## Optimization of operation indexes of contact center considering customer patience and real-time seat conversion

LI Chenglong, LI Junxiang, ZHOU Tingting

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** In order to solve the problem of agent imbalance caused by the traditional contact center serving different types of customers, improve the service quality of the contact center and optimize customer service experience, some multi skilled agents were introduced. Without increasing the number of agents, a part of the multi skilled agents were transferred in real time according to the queue length and agent status. A new contact center model which aimed to minimize customer abandonment was constructed with the goal of the lowest customer abandonment rate. Arena was used to build a model to simulate and analyze the overall performance of the new contact center model and compared with the traditional model to verify its effectiveness. It is found that the new contact center model not only improves the connection rate, but also reduces the waiting time of customers and balances the service intensity of agents.

**Keywords:** contact center; multi service; multi skill; real time transfer; patient value

收稿日期: 2021-04-21

基金项目: 国家自然科学基金资助项目 (72071130, 71871144)

第一作者: 李成龙 (1996-), 男, 硕士研究生. 研究方向: 排队优化. E-mail: 920987273@qq.com

通信作者: 李军祥 (1971-), 男, 教授. 研究方向: 管理科学与工程. E-mail: lijx@usst.edu.cn

随着互联网的发展,许多行业也由传统的电话拨号呼叫中心向服务渠道与服务类型更多的联络中心发展。联络中心处理的服务类型不断增多,但中小型联络中心受到资金有限、雇佣座席数量少,且座席技能单一、不同技能座席相互独立、座席分配根据排班计划较为固定等因素制约。这样的模式灵活性不足:一种服务的座席人员处于高强度的工作状态,另一种服务的座席却常常处于空闲状态,通常会造成某种服务的顾客体验很差以及大量顾客得不到服务。

为提高呼叫中心的服务质量,许多学者对这个问题进行了分析研究。于淼等<sup>[1]</sup>针对带有排队等待信息提示的呼叫中心系统进行了顾客耐心变化随机分布情况下的人力资源配置方法的研究。杨学良等<sup>[2]</sup>利用 Arena 仿真软件对 V 型联络中心与 II 型呼叫中心进行仿真分析对比。于欣<sup>[3]</sup>针对一类具有“前台接入,后台拨出”特征的串联呼叫中心前后台工作内容再分配问题进行了建模研究,探讨了面向再分配结果的优化策略。李军祥等<sup>[4]</sup>通过为会员顾客增设新的座席渠道提出了一种新型的联络中心模型,运用 ProModel 仿真软件对顾客的放弃率进行研究。霍加冕等<sup>[5]</sup>使用线性回归模型分析 IVR 到达量与人工服务到达量之间的关系。Yu 等<sup>[6]</sup>研究了一种既有排队提示等待信息,又有重试选项的呼叫中心模型,证明排队提示等待信息的重要性以及重视重试顾客的重要性。李俊潼等<sup>[7]</sup>研究了一个带有工作休假的 M/M/1 模型,在系统处于不同状态的时候座席具有不同的服务率。刘立芳等<sup>[8]</sup>利用 OPNET 仿真工具对联络中心排队模型的网络性能进行了仿真与安全评估。Jouini 等<sup>[9]</sup>研究了一个客户具有不同优先级、有虚拟延迟信息的多服务台呼叫中心并且提出了一种基于马尔可夫链的方法来估计新到达的虚拟延迟。Dudin 等<sup>[10]</sup>研究了一个具有马尔科夫到达过程、有限和无限缓冲区的多服务器排队系统,用以模拟具有回拨选项的呼叫中心。Legros 等<sup>[11-12]</sup>研究了回拨在不同负荷状态的系统中带来的收益,证明了当大型呼叫中心处于高负荷状态时回拨的重要性和一个渠道是聊天频道的呼叫中心。Dragieva 等<sup>[13]</sup>推导了系统状态稳态分布的计算公式,以及从服务器利用率的角度来表达主要性能宏特性的公式。李嘉兴<sup>[14]</sup>研究了一种可以观测的队列是如何影响服务者对消费者的定价策略的。

何恒<sup>[15]</sup>运用了 Flexsim 仿真软件研究了顾客既有随机到达,又有规律性到达的排队服务系统。张岚等<sup>[16]</sup>对银行排队叫号换汇数据,基于贝叶斯推断对单服务台的马尔科夫与非马尔科夫系统下的 4 个排队模型作参数估计与效果评价。胡修武等<sup>[17]</sup>通过对问题数据特征及优化性质分析提出两个构造性启发式算法,讨论保证员工上班规律性的同班次用工制度对企业人力成本控制的影响。白燕燕等<sup>[18]</sup>研究了一个库存系统,建立了系统的三维 Markov 过程,求出了系统的稳态概率分布。徐秀丽<sup>[19]</sup>研究有灾难到达的可修流体排队中顾客的止步策略,利用 Matlab 进行数值仿真分析,研究了一些系统参数对单位时间内平均社会收益的影响。

## 1 传统呼叫中心与新型联络中心

在传统的研究中,对多服务联络中心的研究基本都是假定各个服务的座席数目固定、每种不同的服务具有不同的优先级或者根据系统情况增加渠道或者增设服务台,往往无法兼顾服务质量与座席成本。本文基于对顾客耐心值的研究,在不增设座席的条件下,引入一定数量的多技能座席,并且考虑在服务时根据队列长度与座席状态对这部分座席进行灵活转移,研究了采取什么样的转移方式(某时刻是否转移座席,转移多少座席)能使联络中心的放弃率最低,这是本文的创新之处。由此建立了新的联络中心模型来平衡各种顾客类型的接通率、座席服务强度与排队等待时长,通过仿真的方法对其进行验证,并与传统的联络中心模型进行仿真对比。

### 1.1 传统呼叫中心

呼叫中心是一种排队模型,包括顾客到达、排队和接受服务 3 个过程。到达以及服务过程均服从一定的概率分布,通常采用的是 FCFS(first come first service)规则。顾客到达呼叫中心会产生队列,在队列中等待时顾客会由于不耐烦选择离去从而产生放弃。队长跟放弃率有很大的关系,其模型如图 1 所示。

### 1.2 新型联络中心

相较于传统的仅使用电话进行服务的呼叫中心,联络中心利用信息技术手段更加注重多种通信渠道(比如电话、QQ、微信、邮件、短信等)的

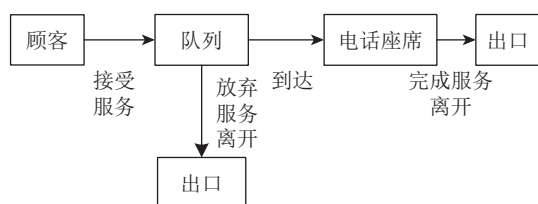


图1 传统呼叫中心模型

Fig.1 Traditional call center model

使用,打造一个可以实现交互式即时沟通的平台。为了解决放弃率高、等待时间长的问题,本文构建了一种新型的联络中心,研究了有顾客不耐烦的情况,并在模型中引入以下措施:

a. 为顾客提供排队等待时间提示。当顾客请求服务,若是服务这种类型顾客的座席全部处于忙碌状态并且顾客愿意等待时,顾客将会收到一个等待时间提示,等待座席有空闲之后对其进行服务。当提示的等待时间超过顾客的耐心值时顾客才会离开。

b. 引入部分多技能座席。单技能座席只能处理一种类型顾客,引入多技能座席能处理更多类

型的顾客。在本文研究两种类型顾客的情况下有3种座席:只能服务第一类顾客的单技能座席、只能服务第二类顾客的单技能座席和两类顾客都能服务的多技能座席,分别将这3类座席称为第Ⅰ,Ⅱ,Ⅲ类座席。

c. 对多技能座席进行灵活的实时转移。一个座席在同一时刻只能处理一个顾客,在新型联络中心模型中,第Ⅲ类座席的分配将会根据队列的长度以及座席的忙闲进行转移。当第一类顾客的等待人数超过阈值时,前一段时间内服务第二类顾客并且此刻闲置的第Ⅲ类座席将会被转移去服务第一类顾客;而为第一类顾客服务的第Ⅲ类座席开始有较多闲置时(达到座席空闲阈值),同样也会将其转移去服务第二类顾客。需要注意的是,只有满足阈值条件(队列阈值或者座席空闲阈值)时第Ⅲ类座席才会转移去服务另一类顾客,否则即使闲置也不会进行转移。这样做是为了保证第Ⅲ类座席在一段时间内能专注服务某种类型的顾客,保证服务质量。其模型如图2所示。

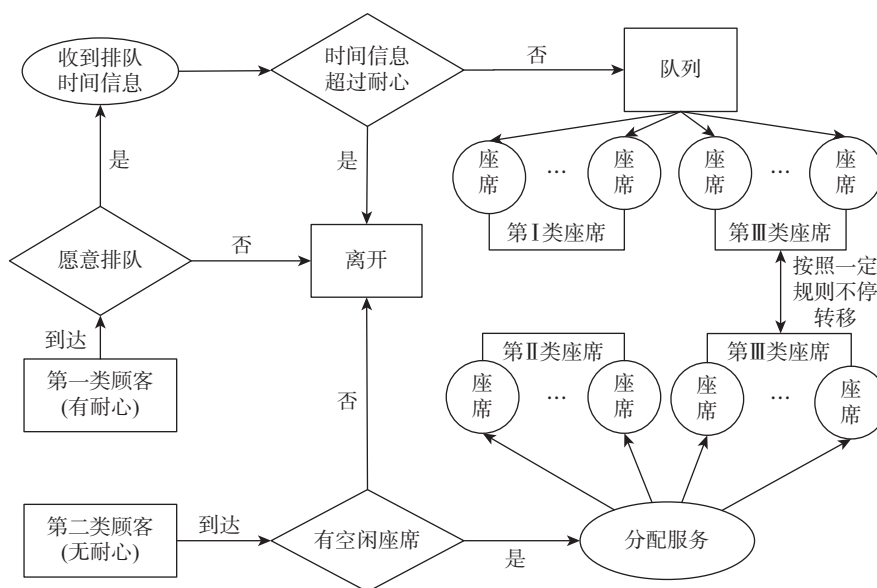


图2 新型联络中心模型图

Fig.2 A new contact center model

## 2 问题描述与模型建立

### 2.1 问题描述

#### 2.1.1 基于队列阈值与空闲座席阈值的座席实时转移

本文将研究一个有实时排队提示信息、座席实时调整的联络中心。考虑了顾客耐心不一样的

情况:两种类型顾客中第一类顾客具有耐心值(座席全部占线则愿意等待),第二类顾客是不耐烦的(一旦座席全部占线则直接离开)。两种类型的顾客到达率(单位时间内平均到达顾客数)分别为 $\lambda_1$ ,  $\lambda_2$ 。第Ⅰ,Ⅱ,Ⅲ类座席的数量分别为 $m_I$ ,  $m_{II}$ ,  $m_{III}$ 。为方便表述规定:无论是哪类座席,只要为第一类顾客服务,就称之为座席1,只要为第

二类顾客服务,就称之为座席2。座席1与座席2的初始人数分别为 $m, n(m > n)$ 。并且规定在系统初始时刻第Ⅲ类座席都为座席2,所以 $m = m_I, n = m_{II} + m_{III}$ 。令 $t$ 时刻座席1与座席2的数目分别为 $M_1(t)$ 与 $M_2(t)$ ,其中 $t \in [0, T], [0, T]$ 是联络中心的服务时段。在第一类顾客到达时,如果同时有第Ⅰ类和第Ⅲ类座席空闲,优先将其分配给第Ⅰ类座席;在第二类顾客到达时,如果同时有第Ⅱ类和第Ⅲ类座席空闲,优先将其分配给第Ⅱ类座席。这保证了第Ⅲ类座席会优先空闲。令在 $t$ 时刻处于忙碌状态的座席1与座席2的数目分别为 $H_1(t)$ 和 $H_2(t)$ ,可得在 $t$ 时刻处于空闲状态的座席1与座席2个数分别为 $I_1(t) = M_1(t) - H_1(t), I_2(t) = M_2(t) - H_2(t)$ 。

本文对第Ⅲ类座席进行了实时的转移,实现的方式如下:用 $N(t)$ 代表在 $t$ 时刻等待的第一类顾客数量, $a$ 是队列阈值, $b$ 是座席1的空闲数阈值。用 $Q_1(t)$ 判断是否在 $t$ 时刻将服务第二类顾客的第Ⅲ类座席转移服务第一类顾客。只有当 $N(t)$ 超过阈值 $a$ 、设定转移的座席数 $m_1(m_1 \leq m_{III})$ 小于座席2的空闲座席数 $I_2(t)$ 时才进行转移。若转移则 $Q_1(t) = 1$ ,否则 $Q_1(t) = 0$ 。另外要求 $m_1 \leq a$ ,否则转移的座席过多会直接产生空闲座席。因此

$$Q_1(t) = \begin{cases} 1, & N(t) > a, \\ & m_1 \leq \min\{I_2(t), m_{III}, a\} \\ 0, & \text{其他} \end{cases} \quad (1)$$

用 $Q_2(t)$ 判断是否在 $t$ 时刻将服务第一类顾客的第Ⅲ类座席转移服务第二类顾客,只有当座席1的空闲座席 $I_1(t)$ 超过阈值 $b$ 且座席2没有空闲时才进行转移,若发生转移则 $Q_2(t) = 1$ ,否则 $Q_2(t) = 0$ 。转移座席个数为 $m_2(m_2 \leq m_{III})$ ,规定 $m_2 \leq I_1(t)$ 来保证转移座席数小于空闲座席数。因此

$$Q_2(t) = \begin{cases} 1, & I_1(t) > b, m_2 \leq \min\{m_{III}, I_1(t)\}, \\ & I_2(t) = 0 \\ 0, & \text{其他} \end{cases} \quad (2)$$

记从 $[0, T]$ 时间段内进入系统的第一类顾客数、第二类顾客数、全部顾客数分别为 $R_1(T), R_2(T), R(T)$ 。而截止到 $t$ 时刻,即在 $[0, t]$ 时段内进入系统的第一类顾客数、第二类顾客数、全部顾客数分别为 $R_1(t), R_2(t), R(t) = R_1(t) + R_2(t)$ ,规定在 $[0, t]$ 时段内第 $c$ 个第一类顾客进入的时刻为 $r_c(c = 1, 2, 3, \dots, R_1(t))$ ,第 $d$ 个第二类顾客进入的时

刻为 $g_d(d = 1, 2, 3, \dots, R_2(t))$ 。 $C(t)$ 表示在 $[0, t]$ 时间段内累计第Ⅲ类座席转移去服务第一类顾客的次数, $B(t)$ 表示在 $[0, t]$ 时间段内累计第Ⅲ类座席转移去服务第二类顾客的次数。那么

$$C(t) = \sum_{i=1}^{R_1(t)} Q_1(r_i) \quad (3)$$

$$B(t) = \sum_{i=1}^{R_2(t)} Q_2(g_i) \quad (4)$$

可得 $M_1(t), M_2(t)$ 两种座席数随时间 $t$ 变化的函数为

$$M_1(t) = m + m_1 C(t) - m_2 B(t) \quad (5)$$

$$M_2(t) = n - m_1 C(t) + m_2 B(t) \quad (6)$$

记在 $[0, T]$ 时间段内座席1与座席2的平均数目分别为 $G_1(T)$ 与 $G_2(T)$ ,则

$$G_1(T) = \frac{\int_0^T M_1(t) dt}{T} \quad (7)$$

$$G_2(T) = \frac{\int_0^T M_2(t) dt}{T} \quad (8)$$

### 2.1.2 座席转移的收益与成本

进行座席转移的目的是让顾客尽早得到服务,顾客可以从座席转移中受益,每次转移也会需要成本。假设在每次转移中每个转移去服务第一类顾客的座席带来的收益是 $k_1$ ,成本是 $h_1$ ,每个转移去服务第二类顾客的座席带来的收益是 $k_2$ ,成本是 $h_2$ 。当 $t = T$ 时, $C(T)$ 与 $B(T)$ 分别代表 $[0, T]$ 时间段内转移的累计次数( $t = T$ ,代入式(3)和式(4))。可以计算得到总收益为 $m_1 k_1 C(T) + m_2 k_2 B(T)$ ,总成本为 $m_1 h_1 C(T) + m_2 h_2 B(T)$ 。为保证转移是有效果的,需要保证收益大于成本,因此有约束

$$m_1 h_1 C(T) + m_2 h_2 B(T) < m_1 k_1 C(T) + m_2 k_2 B(T) \quad (9)$$

### 2.1.3 实时提示排队等待时间

第一类顾客进入时,如果发现有人在排队,会有 $\alpha$ 的概率直接离开系统。顾客以 $1 - \alpha$ 的概率进入队列之后会收到提示等待的时长。等待时长的估计中最常用的是基于队长的等待时间估计,其定义公式为

$$\text{等待时长} = \frac{\text{队列长度} + 1}{\text{当前座席个数} \times \text{单个座席服务率}}$$

设第 $c$ 个第一类顾客提示的等待时长用 $W_c$ 表示,当第 $c$ 个第一类顾客在 $r_c$ 时刻到来时,如果有座席空闲,则提示的等待时长 $W_c$ 为0。如果所有座席均处于忙碌状态,在 $r_c$ 时刻有 $N(r_c)(N(r_c) \leq a)$ 个人在排队,那么这个新到达的顾客要等他前面这

$N(r_c)$ 个人都进入被服务状态且刚好有一个被服务完,他才能进入被服务的状态。这种情况下,将队长和座席个数代入等待时长公式得到等待时间  $W_c$  为

$$W_c = \frac{N(r_c) + 1}{M_1(r_c)\mu_1} \quad (10)$$

而当所有座席均处于忙碌状态且有  $N(r_c)$  ( $N(r_c) > a$ ) 个人在排队时,由于队列超过阈值会发生座席的转移,会有  $m_1$  个第Ⅲ类座席过来服务第一类顾客,在队列中的人有  $m_1$  个立即接受服务,此时队列的人数会变成  $N(r_c) - m_1$ , 这种情况下等待时间  $W_c$  为

$$W_c = \frac{N(r_c) - m_1 + 1}{M_1(r_c)\mu_1} \quad (11)$$

整理得

$$W_c = \begin{cases} 0, & H_1(r_c) < M_1(r_c) \\ \frac{N(r_c) + 1}{M_1(r_c)\mu_1}, & H_1(r_c) = M_1(r_c), N(r_c) \leq a \\ \frac{N(r_c) - m_1 + 1}{M_1(r_c)\mu_1}, & N(r_c) > a \end{cases} \quad (12)$$

令  $\bar{W}$  为  $[0, T]$  时间段内进入的  $R_1(T)$  个第一类顾客接收到提示等待时间的平均值,可以得到

$$\bar{W} = \frac{\sum_{i=1}^{R_1(T)} W_i}{R_1(T)} \quad (13)$$

#### 2.1.4 座席的服务强度

根据文献 [1], 设第一类顾客具有耐心值  $T_{\max}$  (最大愿意等待时间), 且服从参数为  $\theta$  的指数分布。令  $p_1$  为第一类顾客进入队列接收到提示等待时间后的放弃概率。当提示的等待时间超过顾客耐心, 顾客会放弃, 其概率为

$$p_1 = P(T_{\max} < \bar{W}) = 1 - e^{-\theta\bar{W}} \quad (14)$$

$W_1(T)$  为在  $[0, T]$  时间段内座席 1 全部占线的总时长 (即有第一类顾客需要排队的时长),  $W_2(T)$  为  $[0, T]$  时间段内座席 2 全部占线的总时长。可以得到在  $[0, T]$  时间段内第一类顾客排队率为  $p_2 = \frac{W_1(T)}{T}$ , 座席 2 所有座席都占线的时间与总时间的比率为  $p_3 = \frac{W_2(T)}{T}$ 。

设每个座席 1 与座席 2 的服务率 (单位时间内平均服务的顾客数) 分别为  $\mu_1$ ,  $\mu_2$  (假定不同种类座席对同一类顾客的服务率一样)。有效到达率是指在单位时间内到达系统并且成功接受到服务的顾客的个数, 令  $\lambda_1'$ ,  $\lambda_2'$  分别代表第一类、第二类

顾客的有效到达率。 $\lambda_1'$  由单位时间内进入直接接受服务的顾客和在队列中等待没有离开的顾客组成, 有

$$\lambda_1' = (1 - p_2)\lambda_1 + p_2(1 - \alpha)(1 - p_1)\lambda_1 \quad (15)$$

$\lambda_2'$  由单位时间内未放弃的第二类顾客组成, 有

$$\lambda_2' = (1 - p_3)\lambda_2 \quad (16)$$

根据排队论知识, 整个系统的服务率等于系统中的座席个数与单个座席服务率的乘积。根据式 (7) 和式 (8), 在得到了  $[0, T]$  内座席 1 与座席 2 的平均数目与单个座席的服务率之后, 可以计算出在  $[0, T]$  内所有座席 1 与座席 2 的平均服务率  $\mu_1(T)$ ,  $\mu_{II}(T)$  分别为

$$\mu_1(T) = \mu_1 G_1(T) \quad (17)$$

$$\mu_{II}(T) = \mu_2 G_2(T) \quad (18)$$

服务强度是指单个服务台在单位时间内用于为顾客提供服务的平均服务时间, 反映了服务台的繁忙程度, 其计算公式为

$$\text{服务强度} = \frac{\text{有效到达率}}{\text{座席个数} \times \text{座席服务率}}$$

令  $\pi_1$ ,  $\pi_2$  分别代表  $[0, T]$  时间段内单个座席 1 的服务强度和单个座席 2 的服务强度, 有

$$\pi_1 = \frac{\lambda_1'}{\mu_1 G_1(T)} = \frac{\lambda_1'}{\mu_1(T)} \quad (19)$$

$$\pi_2 = \frac{\lambda_2'}{\mu_2 G_2(T)} = \frac{\lambda_2'}{\mu_{II}(T)} \quad (20)$$

两种座席的服务强度应该保持在一个合理的范围内, 既不能太高也不能太低, 设定:  $[0, T]$  时间段内座席 1 的服务强度  $\pi_1$ 、座席 2 的服务强度  $\pi_2$  的下限分别为  $\vartheta_1$ ,  $\vartheta_2$ , 上限分别为  $\beta_1$ ,  $\beta_2$ 。因此得到约束:

$$\vartheta_i \leq \pi_i \leq \beta_i, i = 1, 2 \quad (21)$$

#### 2.1.5 顾客的放弃

令在  $[0, T]$  时间段, 第一类顾客、第二类顾客、整个系统的放弃人数分别为  $D_1(T)$ ,  $D_2(T)$ ,  $D(T)$ 。 $D_1(T)$  主要包括两部分: 一部分是发现有队列直接离开系统的顾客, 另一部分是接收到的提示等待时间大于其耐心值的顾客, 即

$$D_1(T) = R_1(T)p_2\alpha + R_1(T)p_2(1 - \alpha)p_1 \quad (22)$$

$D_2(T)$  是第二类顾客发现座席全部被占线直接离开的人数, 即

$$D_2(T) = R_2(T)p_3 \quad (23)$$

因此可以得到总的放弃人数为

$$D(T) = D_1(T) + D_2(T) \quad (24)$$

故在 $[0, T]$ 时间段内整个系统的放弃率为

$$A(T) = \frac{D(T)}{R(T)} \quad (25)$$

## 2.2 模型建立

本文的目标函数是 $A(T)$ 。结合约束条件, 本文所建立的数学模型如下:

$$\min A(T) = \frac{D(T)}{R(T)} = \frac{\lambda_1 W_1(T) \left[ \alpha + (1 - \alpha) \left( 1 - e^{-\frac{-\theta \sum_{i=1}^{R_1(T)} W_i}{\lambda_1 T}} \right) \right] + \lambda_2 W_2(T)}{R(T)}$$

$$\text{s.t.} \begin{cases} \vartheta_1 \leq \frac{\left( 1 - \frac{W_1(T)}{T} \right) \lambda_1 + \frac{W_1(T)}{T} (1 - \alpha) e^{-\frac{-\theta \sum_{i=1}^{R_1(T)} W_i}{\lambda_1 T}} \lambda_1}{\mu_1 \int_0^T \left( m + m_1 \sum_{i=1}^{R_1(t)} Q_1(r_i) - m_2 \sum_{i=1}^{R_2(t)} Q_2(g_i) \right) dt} \leq \beta_1 \\ \vartheta_2 \leq \frac{\left( 1 - \frac{W_2(T)}{T} \right) \lambda_2}{\mu_2 \int_0^T \left( n - m_1 \sum_{i=1}^{R_1(t)} Q_1(r_i) + m_2 \sum_{i=1}^{R_2(t)} Q_2(g_i) \right) dt} \leq \beta_2 \\ m_1 h_1 \sum_{i=1}^{R_1(T)} Q_1(r_i) + m_2 h_2 \sum_{i=1}^{R_2(T)} Q_2(g_i) < m_1 k_1 \sum_{i=1}^{R_1(T)} Q_1(r_i) + m_2 k_2 \sum_{i=1}^{R_2(T)} Q_2(g_i) \\ Q_1(t), Q_2(t) \in \{0, 1\}, \quad t \in [0, T] \\ m_1 \leq \min\{I_2(t), m_{\text{III}}, a\}, m_2 \leq \min\{m_{\text{III}}, I_1(t)\}, \quad t \in [0, T] \end{cases}$$

其中, 目标函数和约束条件中 $\sum_{i=1}^{R_1(T)} W_i$ 的定义以及展开见式(9)与式(10)。前两个约束条件是座席服务强度的约束, 第三个条件是为了保证座席转移的收益大于成本, 第四个条件规定其是0-1变量, 最后一个条件是座席转移数目的约束。决策变量为 $Q_1(r_c)(c = 1, 2, 3, \dots, R_1(T))$ ,  $Q_2(g_d)(d = 1, 2, 3, \dots, R_2(T))$ ,  $m_1$ 和 $m_2$ 。 $Q_1(r_c)$ ,  $Q_2(g_d)$ 的展开见式(1)与式(2)。模型的含义是: 在保证座席服务强度在合理范围之内、座席转移的收益大于成本的条件下, 如何采取转移策略(某时刻是否要转移第Ⅲ类座席, 转移的数量是多少)才能使顾客的放弃率最低。

## 3 Arena 软件仿真

### 3.1 仿真软件介绍

Arena 是计算机模拟软件包。用户可以根据需要设定参数, 建立仿真模型对实际系统进行分析 and 处理, 找出资源配置的最佳方案。由于建立的模型需要对实时排队状况以及座席实时安排情况作出调整, 因此本文选择其进行联络中心的建模研究, 并将本文提出的新型模型与有多技能座席但不采取实时转移的联络中心模型(下文称之为模

型二)和传统联络中心模型(单技能座席)进行比较。

### 3.2 仿真模块分析

本文主要用到了 Arena 仿真软件中的 Entity, Create, Resource, Process, Decide, Alter, Assign, Optquest。

#### a. Entity。

Entity 负责定义请求服务的对象: 顾客。主要定义两种 Entity: 第一类顾客与第二类顾客。

#### b. Create。

Create 负责创建好实体在模型中的到达过程。假设第一类顾客到达率服从 $\lambda_1 = 1.25$ 的泊松分布, 第二类顾客到达率服从 $\lambda_2 = 0.55$ 的泊松分布, 仿真时间为 $T = 600$  min。在 $[0, 600]$ 时间段内共有 $R(600) = 1000$ 个顾客进入系统, 第一类顾客 $R_1(600) = 700$ 个, 第二类顾客 $R_2(600) = 300$ 个。

#### c. Resource。

Resource 主要用于分配各种类型初始座席个数, 设置座席1与座席2个数分别为 $m = 20$ 与 $n = 12$ 。

其中, 第Ⅰ类座席数 $m_I = 20$ , 第Ⅱ类座席数 $m_{II} = 9$ , 第Ⅲ类座席数 $m_{III} = 3$ 。

#### d. Process。

在 Process 模块中实现座席的服务过程, 此模



块涉及到座席的服务效率以及消耗资源个数的设置。座席1的服务效率服从 $\mu_1 = 1/20$ 的指数分布,座席2的服务效率服从 $\mu_2 = 1/10$ 的指数分布,每次服务消耗的资源个数为1。在Action中需要选择Seize Delay Release,这样每次完成处理过程之后将会释放此资源以便其继续处理后续的服务。

#### e. Decide。

Decide根据系统运行的当前状态进行判定,决定运行的下一个状态需要进行什么调整,座席的转移、顾客的离开等行为需要这个模块的加入。用户的耐心值设定是本文的核心任务,需要对其给出一个合理的值,根据客户服务技术提供商Fast Customer<sup>[20]</sup>的报告称顾客的平均等待时间跟其所接受服务的时间相关,一般在服务时间

的20%~80%之间,再参考文献[1]中的参数设置,令顾客的耐心值 $T_{\max}$ 服从 $\theta = 0.5$ 的指数分布。 $((NQ(\text{Custom1.Queue}) + 1))/(\text{MR}(\text{Tech 1}) * \text{EXPO}(20))$ 是提示等待时间的表达式。其中 $\text{MR}(\text{Tech 1})$ 代表当前座席1的个数, $NQ(\text{Customer1.Queue})$ 代表第一类顾客队列长度,需要注意的是,在仿真中队列的长度实时就能调整好,所以不需要像数学模型中那样分类讨论。设队列阈值 $a = 5$ ,空闲座席2阈值 $b = 3$ 。

#### f. Alter。

Alter可实现座席转移,在经过Decide模块之后,若需要进行座席转移便会对座席的个数进行调整,具体转移规则见表1。初始规定 $m_1 = 3$ ,  $m_2 = 3$ , Tech 2代表座席2, Tech 3代表第Ⅲ类座席。

表1 座席转移规则

Tab.1 Transfer rules of agents

转移方式	Decide模块需要满足的条件	表达式
第Ⅲ类座席转移服务第一类顾客	$(\text{MR}(\text{Tech 2}) - \text{NR}(\text{Tech 2})) > m_1 \&\& NQ(\text{QVX Service.Queue}) \geq a$	$\text{Tech 1,} + m_1$ $\text{Tech 3,} - m_1$
第Ⅲ类座席转移服务第二类顾客	$\text{MR}(\text{Tech 1}) = \text{NR}(\text{Tech 1}) \&\& (\text{MR}(\text{Tech 1}) - \text{NR}(\text{Tech 1})) > m_2$	$\text{Tech 3,} - m_2$ $\text{Tech 2,} + m_2$

#### g. Assign。

Assign模块用于定义一些仿真模块中不自带的变量,本文定义座席转移的收益、成本与座席转移次数。每当一个实体经过Alter就记录一次转移次数,在模块中两种转移总次数分别命名为Time1与Time2,并且在模块中定义 $k_1$ 服从参数为5的正态分布, $h_1$ 服从参数为3的正态分布, $k_2$ 服从参数为2的正态分布, $h_2$ 服从参数为3的正态分布。

#### h. Optquest。

Optquest模块用来实现约束。座席的服务强度是仿真系统自定义出来的,座席1,2的服务强度表达式为 $\text{Tech 1.Utilization}$ ,  $\text{Tech 2.Utilization}$ 。令服务强度下限 $\vartheta_1 = 0.3$ ,  $\vartheta_2 = 0.3$ ,上限分别为 $\beta_1 = 0.8$ ,  $\beta_2 = 0.8$ 。在Optquest模块中约束: $\text{Time1} * k_1 + \text{Time2} * k_2 > \text{Time1} * h_1 + \text{Time2} * h_2$ ,  $0.3 < \text{Tech1.Utilization} <$

$0.8$ ,  $0.3 < \text{Tech 2.Utilization} < 0.8$ 。

### 3.3 仿真结果分析

在所有条件相同的情况下对3种模型:新型模型、有多技能座席但不采取实时转移的联络中心模型(模型二)和传统联络中心模型(单技能座席)进行对比分析。每次运行的时长为600 min,次数为10次,具体的参数设置仿真条件如表2所示。

因为接通率与放弃率之和为1,所以放弃率越低,接通率越高。接通率越高越好,放弃率越低越好。在仿真结果中本文选取接通率进行分析,最终得到的仿真结果数据如表3所示。

从表3中可以看出:在总计1000人的到达人次中,在总接通率上新模型达到了94%,并且两种类型顾客的接通率都很高,也都在95%左右,顾客平均等待时间是1.2 min。模型二总接通率为

表2 相关仿真参数设置

Tab.2 Related simulation parameter setting

模型类别	第一类顾客到达率/ (个 $\cdot\text{min}^{-1}$ )	第二类顾客到达率/ (个 $\cdot\text{min}^{-1}$ )	座席1服务率/ (个 $\cdot\text{min}^{-1}$ )	座席2服务率/ (个 $\cdot\text{min}^{-1}$ )	座席1 个数/个	座席2 个数/个
新模型	1.25	0.55	1/20	1/10	20	12
模型二	1.25	0.55	1/20	1/10	20	12
传统模型	1.25	0.55	1/20	1/10	20	12

表 3 仿真结果  
Tab.3 Simulation results

模型类别	到达顾客总数/个, 第一类顾客总数/个, 第二类顾客总数/个	总接通率, 第一类顾客接通率, 第二类顾客接通率	顾客平均等待时间/min	座席1服务强度, 座席2服务强度
新模型	1000, 700, 300	0.94, 0.95, 0.94	1.2	0.84, 0.41
模型二	1000, 700, 300	0.88, 0.90, 0.78	2.1	0.74, 0.60
传统模型	1000, 700, 300	0.83, 0.79, 0.99	10.2	0.85, 0.37

89%, 第一类顾客的接通率达到了 90%, 但是第二类顾客的接通率只有 78%, 顾客平均等待时间为 2.1 min。传统模型总接通率最低只有 83%, 第一类顾客接通率也是最低的 79%, 第二类顾客接通率达到了 99%, 虽然基本没有放弃的, 但排队等待时间高达 10.2 min。

图 3 是 3 种模型放弃人数与队列长度随仿真时间变化的对比。传统模型的队列长度随着仿真时间变化始终维持在较高水平, 最多的时候有超过 20 人在排队。这意味着在传统模型中整个系统的状态很拥堵, 顾客的等待时间很长, 服务质量很差。所以随着时间变化它的放弃人数也是增长最快的。模型二队列长度有了减少, 放弃人数也下降了, 但是放弃人数的增长速度也很快。而新模型的队列长度随着时间变化在比较小的范围内波动, 队列长度始终维持在不到 5 人, 远远短于传统的模型, 放弃人数也上升得很缓慢。引入多技能座席并进行实时转移, 对队列长度和放弃人数上的优化效果很显著。

图 4 是 3 种模型随着顾客耐心阈值变化接通率与等待人数的变化情况。可以看出: 在顾客越来越有耐心的情况下 3 种模型的接通率都会增加, 而传统模型的接通率对耐心阈值的敏感程度大于其他两种模型, 并且在顾客特别有耐心的情况下接通率是最高的。这是由于顾客愿意等待, 所以放弃较少, 故传统模型等待队列中的人数也是上升最快的, 相较于其他两种模型, 传统的联络中心模型容易受顾客耐心值的影响, 在面对不耐烦顾客时这种模型的接通率就比较低, 其他的两种模型则能很好地适应各种耐心值的顾客。新模型在保证了接通率的同时, 也降低了顾客的等待时间。

图 5 是新模型在不同的顾客耐心值与队列阈值下接通率的变化。随着顾客耐心值的提升与队列阈值的下降, 新模型总体的接通率是上升的。不过, 对于队列阈值比较短(意味着进行座席转移更加频繁)的模型而言, 它的接通率相对来说受

顾客耐心值的影响更小一点, 也拥有更加高的接通率。

接下来要分析的是初始的座席分配策略对这 3 种模型座席服务强度、排队等待时长以及接通率的影响。假设总的座席人员的数目是不变的, 而初始座席数目对 3 种模型也会产生一些影响。对于新模型与模型二来说, 这里的座席 1 初始人数是第 I 类座席与第 III 类座席之和, 传统模型只有第 I 类座席。

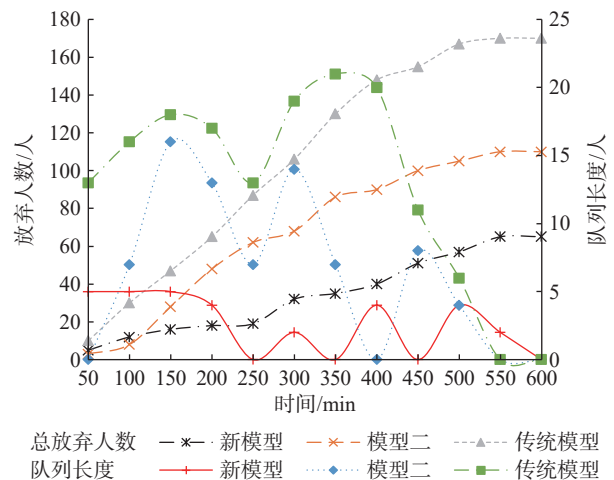


图 3 放弃人数与队列长度随时间变化的对比图

Fig. 3 Comparison of queue length and number of abandonment with time

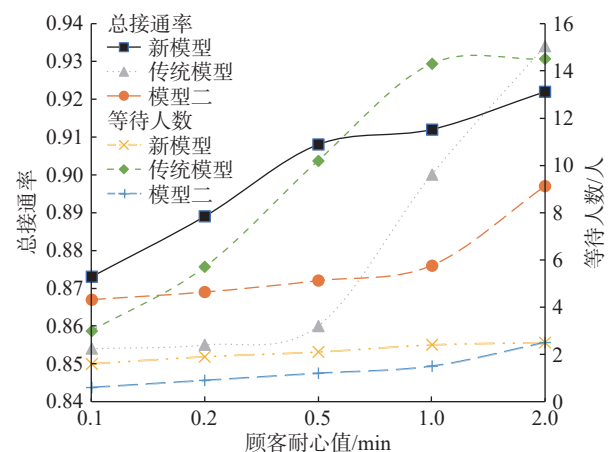


图 4 总接通率与等待人数随顾客耐心值变化的对比图

Fig. 4 Comparison of total connection rate and the waiting number of three models with customer patience value



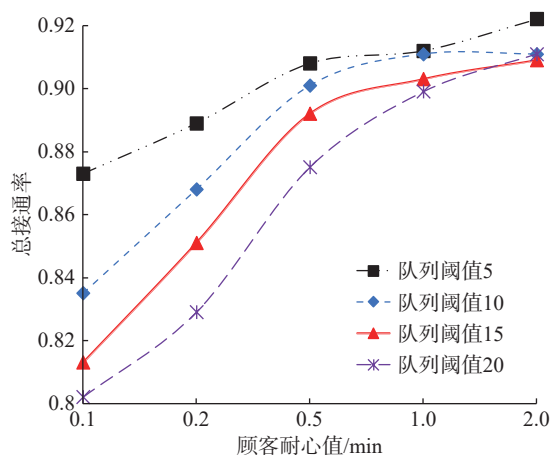


图5 基于队列阈值下总接通率与顾客耐心值之间的关系

Fig. 5 Relationship between total call rate and customer patience based on queue threshold

图6是两种座席的服务强度随座席1的初始人数变化的关系对比图。可以看出:无论在什么样的初始座席分配条件下,新模型两种座席的服务强度都始终维持在一个比较稳定的状态,这是因为座席互相转移策略实施中,第一类顾客处于拥堵状态时,会将处于空闲状态的座席转移服务第一类顾客,而一旦其又闲置时,会再转移回去,这种策略使得两种座席的服务强度受座席1初始分配数量的影响小。模型二座席服务强度的稳定性相对较差。传统模型两种座席的服务强度随初始座席个数的变化波动很大,服务强度受座席人数影响,十分不稳定。

图7和图8分别展示了3种模型两类顾客的接通率、总接通率和等待时间随座席1初始人数的

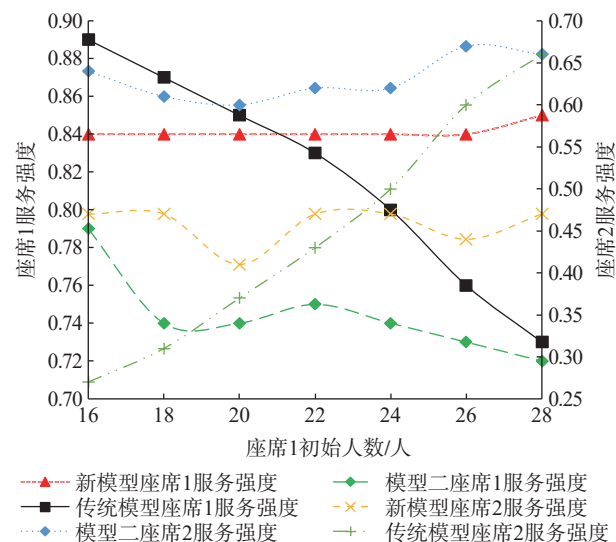


图6 座席1和2服务强度随座席1初始人数变化的对比

Fig. 6 Comparison of service intensity of agent 1 and agent 2 with the initial number of agent 1

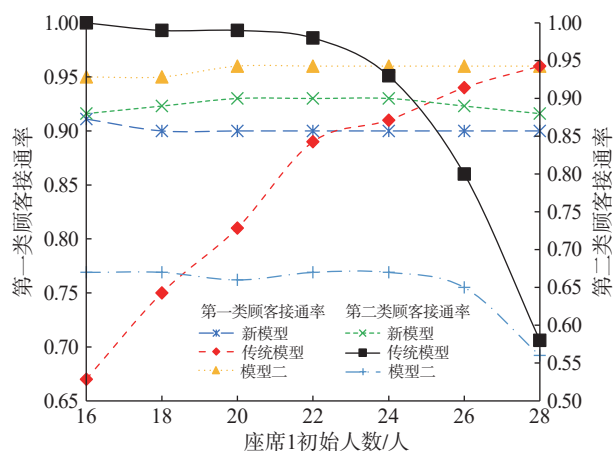


图7 第一、二类顾客接通率随座席1初始人数变化对比

Fig. 7 Comparison of the change of the first and second class customer connection rate with the initial number of agent 1

变化。最稳定的仍然是新模型,各个接通率都在90%左右。模型二随着座席1人数的增加虽然第一类顾客接通率维持在95%左右,但是第二类顾客的接通率却达不到70%。传统模型的各个接通率很不稳定,一旦座席个数分配不好,整个模型的接通率将会很不理想。不论在什么样的座席分配情况下,新模型与模型二都能保证顾客在5 min之内接受到服务且新模型的稳定性更高。

综上,通过在同一条件下的仿真发现,对比传统的联络中心与模型二,新型联络中心在提高总体接通率的同时大大减少了顾客排队等待时间,让顾客的等待时间从10 min下降到了2 min,也在一定程度上平衡了座席的服务强度。模型二相对传统模型有了改进,虽然维持了较高的第一类顾客的接通率,但是却牺牲了第二类顾客的接通率。而传统的联络中心模型不具有灵活性,初

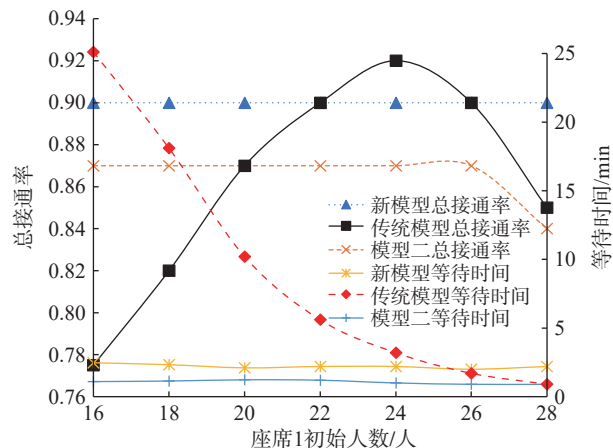


图8 总接通率与等待时间随座席1初始人数变化的对比

Fig. 8 Comparison of the total connection rate and waiting time with the initial number of agent 1 in three models

始座席数量的分配对整个联络中心的影响是决定性的,也就是说联络中心的各项指标:座席的服务强度、排队的等待时间和接通率对座席数量的分配都很敏感。引入多技能座席并且将其实施转移的新联络中心模型相较于其他两种模型,拥有更好的稳定性,能更好地适应各种座席分配的情况,并且同时优化了顾客的接通率和排队等待时长。

## 4 结 论

针对传统单技能座席联络中心座席分配上的不足,提出了一种引入多技能座席并将其进行实施转移的新型联络中心模型。研究发现,在顾客耐心值以及座席分配变化的条件下,根据实时队列长度、座席状态和数量对多技能座席进行转移能够有效优化联络中心的接通率、座席服务强度、顾客等待时间等运营指标。新模型相较于其他两种模型在各项运营指标上都有较大的提升,并且受顾客耐心值、座席分配等因素的影响比较小,稳定性较强,处理各个渠道的任务量时更加平衡。其优化结果给中小型联络中心如何安排座席、提高效率提供了科学依据。本文使用了Arena仿真软件,对模型进行了不同参数下的多次仿真运行,结果较为可靠。

在目前的研究中仍有可以方面,例如可以引入更多的优化指标、考虑除了不耐烦行为外更多的行为要素(例如等待过程中的满意度等),未来将会对这些方向展开研究。

### 参考文献:

- [1] 于森, 宫俊, 雒兴刚, 等. 考虑顾客耐心的呼叫中心人力资源配置模型[J]. 系统工程学报, 2013, 28(5): 686–693.
- [2] 杨学良, 李军祥, 台玉红. 基于Arena仿真的联络中心运营效率对比研究[J]. 系统仿真学报, 2014, 26(11): 2739–2744.
- [3] 于欣, 李娜. 基于马尔可夫的串联呼叫中心系统建模与工作分配策略研究[J]. 工业工程与管理, 2017, 22(1): 65–72.
- [4] 李军祥, 藏万斌. 基于耐心阈值下联络中心放弃率的仿真研究[J]. 系统仿真学报, 2021, 33(1): 169–179.
- [5] 霍加冕, 谢金贵. 呼叫中心到达过程的建模与预测[J]. 中国科学技术大学学报, 2017, 47(9): 770–777.
- [6] YU M, TANG J F, KONG F W, et al. Fluid models for call centers with delay announcement and retrials[J]. Knowledge-Based Systems, 2018, 149: 99–109.
- [7] 李俊潼, 李涛, 徐金萍. 带有反馈的M/M/1重试工作休假排队模型[J]. 应用数学进展, 2019, 8(2): 210–218.
- [8] 刘立芳, 赵丽莎, 齐小刚, 等. 基于排队模型的网络性能仿真与安全评估技术[J]. 系统仿真学报, 2018, 30(1): 191–196.
- [9] JOUINI O, DALLERY Y, AKŞIN Z. Queueing models for full-flexible multi-class call centers with real-time anticipated delays[J]. International Journal of Production Economics, 2009, 120(2): 389–399.
- [10] DUDIN A, KIM C, DUDINA O, et al. Multi-server queueing system with a generalized phase-type service time distribution as a model of call center with a call-back option[J]. Annals of Operations Research, 2016, 239(2): 401–428.
- [11] LEGROS B, JOUINI O, KOOLE G. Optimal scheduling in call centers with a callback option[J]. Performance Evaluation, 2016, 95: 1–40.
- [12] LEGROS B, JOUINI O. On the scheduling of operations in a chat contact center[J]. European Journal of Operational Research, 2019, 274(1): 303–316.
- [13] DRAGIEVA V, PHUNG-DUC T. Two-way communication M/M/1/N retrial queue[C]//Proceedings of the 24th International Conference on Analytical and Stochastic Modeling Techniques and Applications. Newcastle-upon-Tyne: Springer, 2017: 81–94.
- [14] 李嘉兴, 李军. 可观测M/M/1排队系统的高低定价策略[J]. 系统工程, 2020, 38(3): 10–16.
- [15] 何恒, 张顶立, 李惠平, 等. 服务系统顾客到达的非齐次泊松过程仿真研究[J]. 系统工程, 2019, 37(6): 148–154.
- [16] 张岚, 钱夕元. 基于近似贝叶斯计算方法的排队模型参数估计[J]. 上海理工大学学报, 2020, 42(2): 108–114, 121.
- [17] 胡修武, 徐悦, 王秀丽. 呼叫中心坐席人员排班问题优化模型与算法研究[J]. 运筹与管理, 2021, 30(8): 44–51.
- [18] 白燕燕, 岳德权. 服务员休假的库存系统的性能分析与优化[J]. 数学的实践与认识, 2021, 51(18): 142–148.
- [19] 徐秀丽, 刘锦平. 可见情形下有灾难到达的流体可修排队均衡分析[J]. 系统科学与数学, 2021, 41(7): 1972–1984.
- [20] 李大川, 忻展红. 大型呼叫中心人工呼入量的最小二乘支持向量机模型[J]. 控制理论与应用, 2009, 26(7): 815–818.

(编辑: 丁红艺)