



## 基于机器学习的线上线下联合服务模式医生排班算法

张越, 王子翔, 周博, 刘冉, 杨之涛

引用本文:

张越, 王子翔, 周博, 等. 基于机器学习的线上线下联合服务模式医生排班算法[J]. *智能系统学报*, 2025, 20(4): 800–812.

ZHANG Yue, WANG Zixiang, ZHOU Bo, et al. Algorithms for physician scheduling under the online and offline combined service mode based on machine learning[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(4): 800–812.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202404032>

## 您可能感兴趣的其他文章

### 智能化医养融合服务平台关键技术及应用研究

Research on key technologies and applications of intelligent medical and care integration service platform  
*智能系统学报*. 2021, 16(5): 972–988 <https://dx.doi.org/10.11992/tis.202103022>

### 面向机器学习的分布式并行计算关键技术及应用

Key technologies and applications of distributed parallel computing for machine learning  
*智能系统学报*. 2021, 16(5): 919–930 <https://dx.doi.org/10.11992/tis.202108010>

### 多约束下多无人机的任务规划研究综述

A survey of mission planning on UAVs systems based on multiple constraints  
*智能系统学报*. 2020, 15(2): 204–217 <https://dx.doi.org/10.11992/tis.201811018>

### 基于深度学习的椎间孔狭窄自动多分级研究

Deep learning based automatic multi-classification algorithm for intervertebral foraminal stenosis  
*智能系统学报*. 2019, 14(4): 708–715 <https://dx.doi.org/10.11992/tis.201806015>

### 大数据背景下高校招生策略预测

The strategy of college enrollment predicted with big data  
*智能系统学报*. 2019, 14(2): 323–329 <https://dx.doi.org/10.11992/tis.201709011>

### 面对智能导诊的个性化推荐算法

A personalized recommendation algorithm for intelligent guidance  
*智能系统学报*. 2018, 13(3): 352–358 <https://dx.doi.org/10.11992/tis.201711036>

DOI: 10.11992/tis.202404032

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250220.1826.002>

# 基于机器学习的线上线下联合服务模式下 医生排班算法

张越<sup>1</sup>, 王子翔<sup>2</sup>, 周博<sup>1</sup>, 刘冉<sup>1</sup>, 杨之涛<sup>3</sup>

(1. 上海交通大学 工业工程与管理系, 上海 200240; 2. 杭州师范大学 阿里巴巴商学院, 浙江 杭州 311121; 3. 上海交通大学医学院附属瑞金医院 急诊科, 上海 200025)

**摘要:** 线上线下联合的医疗服务模式已经成为我国大型医院普遍采用的新型医疗服务模式, 为了优化大型医院在此类模式下的医生资源配置, 本文研究考虑切换成本的医生排班问题。针对此问题, 建立考虑服务水平限制的医生排班马尔可夫决策过程模型, 并设计近似动态规划算法对马尔可夫决策过程高效求解。进一步, 考虑患者高度时变到达以及医疗服务时长等多维不确定性, 基于合作医院的实际数据, 构建数据驱动的循环神经网络模型, 提出基于数据驱动的线上线下患者排队系统的性能评估方法。数值实验显示, 所提出的方法能够降低医生总工作时长, 并有效控制患者等待时间, 保证系统的高效运行。本文研究结果可为大型医院合理配置线上线下医疗资源提供理论依据和决策支持。

**关键词:** 线上医疗; 医生排班; 时变排队系统; 数据驱动; 深度学习; 马尔可夫决策过程; 近似动态规划; 启发式算法

中图分类号: TP18 文献标志码: A 文章编号: 1673-4785(2025)04-0800-13

中文引用格式: 张越, 王子翔, 周博, 等. 基于机器学习的线上线下联合服务模式下医生排班算法 [J]. 智能系统学报, 2025, 20(4): 800-812.

英文引用格式: ZHANG Yue, WANG Zixiang, ZHOU Bo, et al. Algorithms for physician scheduling under the online and offline combined service mode based on machine learning [J]. CAAI transactions on intelligent systems, 2025, 20(4): 800-812.

## Algorithms for physician scheduling under the online and offline combined service mode based on machine learning

ZHANG Yue<sup>1</sup>, WANG Zixiang<sup>2</sup>, ZHOU Bo<sup>1</sup>, LIU Ran<sup>1</sup>, YANG Zhitao<sup>3</sup>

(1. Department of Industrial Engineering and Management, Shanghai Jiao Tong University, Shanghai 200240, China; 2. Alibaba Business School, Hangzhou Normal University, Hangzhou 311121, China; 3. Department of Emergency, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China)

**Abstract:** The online and offline combined medical service mode has become a new medical service mode generally adopted by large hospitals in China. Under this mode, large hospitals need to allocate physicians to online and offline services, and arrange online and offline scheduling plans for physicians while considering the switching of physicians between the two services. To address this problem, a Markov decision process model for physician scheduling with service level constraints was developed and an approximate dynamic programming algorithm was designed to solve the Markov decision process with high efficiency. Furthermore, considering multi-dimensional uncertainties such as highly time-varying patient arrival and service hours, a data-driven recurrent neural network was constructed based on the real-life data of the cooperative hospital as a performance evaluation method for the online and offline queueing systems. Numerical experiments show that the proposed methods can reduce the total working hours of physicians, effectively control the waiting time of patients, and ensure the high-efficiency operation of the system.

**Keywords:** telemedicine; physician scheduling; time-varying queueing system; data-driven; deep learning; Markov decision process; approximate dynamic programming; heuristic

收稿日期: 2024-04-24. 网络出版日期: 2025-02-21.

基金项目: 国家自然科学基金项目 (72371161); 上海申康医院发展中心管理研究项目 (2024SKMR-19); 上海交通大学中国医院发展研究院研究项目 (CHDI-2024-A-04).

通信作者: 刘冉. E-mail: [liuran2009@sjtu.edu.cn](mailto:liuran2009@sjtu.edu.cn).

近年来, 伴随互联网的发展, 在线医疗服务在全世界范围内蓬勃发展, 形成与传统线下就诊并行的医疗服务方式<sup>[1-2]</sup>。在线医疗是借助无线互联等技术, 通过患者与医生间的实时通信, 完

成远程医疗服务<sup>[3]</sup>。在线医疗服务使医患双方均得到收益。患者通过在线医疗能够降低交通成本,减少在医院的排队时间,降低感染风险;医院则可以辐射更大的患者群体,缓解线下医疗资源挤兑<sup>[4]</sup>。截至2022年我国具有2700多家互联网医院,开展在线诊疗服务超过2590万人次<sup>[5]</sup>,其中约85%为依托实体医院构建的线上医疗平台。相对应地,大量医院形成了线上线下联合的医疗服务模式,同时为患者提供在线和传统线下的医疗服务<sup>[6]</sup>,即科室的部分医生被安排到传统线下诊室工作,部分医生在特定时段提供线上诊疗服务。在这样的服务模式,医院面临着一个重要难题,即如何有效确定每位医生的工作班次,以在不同时段内将有限的医生资源合理分配到线上和线下诊室。然而,医生的统筹排班对医院而言并不容易,这是因为同时考虑线上线下联合医疗服务模式下的医生排班问题具有特定的难点。

首先,线上和线下医疗服务系统都具有很强的随机性和时变性。尽管很多医院推行门诊预约制,但由于患者的到达存在不守时、爽约以及部分患者不经预约直接挂号就诊等现象,线上和线下服务系统的患者到达具有较强的随机性。伴随随机性特征,患者到达还表现出高度时变性,存在显著的患者到达高峰期和低谷期。在如此复杂的随机时变系统中,优化医生线上线下排班以提升服务水平、提升系统运行效率非常困难。进一步,线上、线下服务过程存在显著的差异。线下问诊普遍采用医患之间“一对一”的方式,而典型的线上问诊中医生患者之间采用“一对多”的方式,即一名医生同时为多位患者服务<sup>[7]</sup>。线上问诊中,医生在多个线上患者的交流界面内切换,同时服务多位患者。这导致线上线下患者感受等有显著的差别。如何结合2种服务方式的不同特点,保证服务的公平与质量,成为线上线下医生服务调度难点。最后,高效的排班需要部分医生在线上、线下间进行动态切换,这导致医生调度比传统仅考虑线下的问题更为困难。针对以上难点,本文对线上线下联合服务模式下的医生资源进行排班优化,以提高患者就诊体验,降低医生工作强度。

目前,医生排班问题的研究主要面向线下医疗场景<sup>[8-9]</sup>。Brunner等<sup>[10]</sup>对麻醉科医生柔性排班方案展开优化,决策每位医生的班次开始时间、持续时间以及是否加班,并设计分支定价算法进行求解。Stolletz等<sup>[11]</sup>设计了一种集合覆盖模型,该方法预先确定所有可行的班次类型,然后将医

生分配到各个班次中,以匹配各时段波动的医生需求。Wang等<sup>[12]</sup>针对带回流的急诊服务系统建立离散时间解析模型近似患者等待时间,采用禁忌搜索算法求解相关医生排班问题。Sinreich等<sup>[13]</sup>通过2种迭代启发式方法与仿真相结合,以最小化患者平均停留时间为目标,求解急诊科室内的最优医生工作班次。Liu等<sup>[14]</sup>针对患者随机到达的情况,利用稳态排队模型近似的方法计算患者平均等待时间,设计变邻域搜索算法求解医生排班问题。Zaerpour等<sup>[15]</sup>考虑医生服务新患者速率随其服务时长而递减的特征,以最小化患者总等待时间为目标,建立医生排班的随机规划模型,使用样本均值近似和L-shaped算法求解。祖漫漫<sup>[16]</sup>基于时间序列模型进行患者需求预测,将其转换为医生能力需求约束,在此基础上建立多目标的医生排班模型,采用分层序列法求解。王铖恺等<sup>[17]</sup>考虑新冠疫情期间发热门诊的特殊情况,利用逐点稳态流近似对排队系统建模,结合Benders分解与列生成算法高效求解医生排班问题。兰绍雯<sup>[18]</sup>针对医生资质与服务类型的匹配关系,提出了计划与调配协同的变邻域搜索算法,确定医生人员配置计划表。目前针对线上医生排班问题的研究较少。相关研究集中在探讨患者如何选择线上或线下服务<sup>[6,19]</sup>、研究在线分诊对患者的引导作用<sup>[20]</sup>以及线上和线下医疗服务质量对患者就医决策的影响<sup>[21]</sup>等方面。面向线上线下联合医疗服务模式下的相关问题尚处于起步阶段。除医生排班问题外,如何合理配置资源匹配时变需求同样也是教务排课问题所关注的内容。Hossain等<sup>[22]</sup>考虑课程与教室、教师、学生等多元匹配关系,设计基于粒子群的优化算法最大化教务排课满意度。宋婷等<sup>[23]</sup>通过多类分类器根据排课问题特征对排课问题进行分类,根据分类设计相应迭代局部搜索算法求解。

本文针对该新型服务模式下的医生排班问题进行研究,建立了综合考虑患者平均停留/等待时间、医生工作时间以及加班时间的马尔可夫决策过程模型,利用近似动态规划算法求解排班方案;进一步,提出了基于数据驱动的循环神经网络模型以评估服务系统性能。研究对于提升诊疗效率、缓解医生的工作负荷,具有现实意义。

## 1 医生排班问题

本文聚焦线上线下联合医疗服务模式,研究如何将有限的医生分配到线上、线下诊室中,并对线上线下服务的医生进行工作排班,为到达患

者提供服务。该问题以最小化同科室内所有医生总工作时间为目标,面向各时段的患者等待时间、医生工作规则等约束,决策每位医生在计划周期内的工作班次数量、每个班次的具体工作时间以及所服务患者类型。依据本研究在上海合作医院的调研情况和实际数据,对所研究的问题定义如下文所述。

### 1.1 服务过程假设

本文研究一天的医生排班问题,一天的工作时间被划分为 $|T|$ 个工作时段,每个时段的长度为 $\Delta$ (例如 0.5 h 或者 1 h),一天的全部工作时段用集合 $T=\{1, 2, \dots, |T|\}$ 表示。研究考虑线上和线下诊室中患者候诊和服务的不同,如图 1 所示。线下门诊中,患者随机到达,挂号后在候诊区排队等待接受医生服务,患者遵循先到先服务的规则接受服务。当一名医生空闲,则队列最前的患者被此医生开始一对一的医疗问诊服务,完成后离开系统。设定每个时段 $t$ 内的线下患者到达服从参数为 $\gamma_t^1$ 的泊松过程,每位患者的服务时间服从参数为 $\mu_1$ 的指数分布,即对线下患者排队就诊通过 $M/M/c_t$ 排队模型加以建模<sup>[14]</sup>。在线上问诊中,患者首先通过手机、网页等挂线上就诊号,进入系统在线等待队列,当有医生可为其服务则进入该医生在线诊室,进行在线服务,完成服务后离开系统<sup>[24]</sup>。设定每个时段 $t$ 内的在线患者到达服从参数为 $\gamma_t^2$ 的泊松过程。需要注意的是和线下服务不同,由于线上问诊难以完全实时交流,患者的响应可能存在滞后现象,线上医生和患者之间的问诊并非一对一模式,而是一般采用一对多的模式,即一名医生的线上诊室中可以同时有多名患者<sup>[7]</sup>。当一名患者准备相关材料(例如上传病例、图片等),医生可以和线上诊室的其他患者交流以提供服务,从而提升医生工作时间的利用率。但是一般线上诊室会设定最大“同时服务”的患者数目,以保证服务的质量。本文设定一名医生最多同时为 $K$ 名在线患者提供服务<sup>[25]</sup>。显然,一名医生对每位患者的服务速率受到其同时服务患者的影响,依据调研和已有文献研究<sup>[3]</sup>,本文设定每位线上患者的服务时间服从参数为 $\mu_2(k)$ 的指数分布,其中 $k$ 表示医生同时服务的患者总数。出于表达方便,下文将线上医生“同时服务 $k$ 个患者”称为此医生“处于 $k$ 层级”。线上诊室中采取“层级最低优先”的原则分配医生,即优先将等待队列队首患者分配给层级最低的医生进行服务,这是因为分配前该医生正在服务的患者数最少,负担最轻。如果每位医生都处于最大阈值 $K$ 级,

则新到达的患者处于等待队列中,不能进入任何一个医生的线上诊室<sup>[3]</sup>。

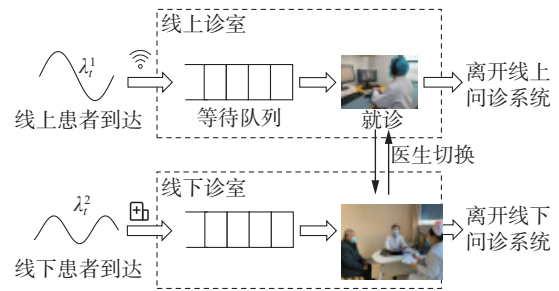


图 1 线上与线下诊室排队系统示意

Fig. 1 Schematic diagram of queuing system in the on-line and offline clinic

### 1.2 医生工作规则

在一天的工作时间内,医生被安排在线上或者线下对患者提供服务。每位医生每天的工作日程由一个或者多个班次组成。一个班次是一段连续工作的时间,班次的开始与结束时间均可以灵活调整。医生上下班工作需要遵守以下规则:

- 1) 各时段线上、线下都至少有一名医生上班。
- 2) 一个医生每天最多安排一个线上诊室的班次与一个线下诊室的班次。
- 3) 一个医生在诊室 $j$ 中的班次长度至少为 $d_j^{\min}$ 个时段,至多为 $d_j^{\max}$ 个时段, $j=1$ 和 $j=2$ 分别表示线上和线下诊室。
- 4) 一个医生在 2 个班次之间的最短休息时间为 1 个时段。
- 5) 医生上班、下班都只能在一个时段的开始时发生。
- 6) 在每天的最后一个时段,工作的医生需要服务完所有患者后才能下班,因此可能产生加班的工作时间。

本文研究的问题目标是在保证服务水平的前提下降低医院的运营成本。保证医疗服务水平显然是必要的,但是由于线上、线下服务模式的差别,其服务水平的衡量方式有显著区别。对线下医疗服务,一般采用线下患者的“平均等待时间”体现服务水平;对线上医疗服务,服务水平指标则为线上患者的“平均停留时间”,即患者从进入到离开线上排队系统的总共时间的均值<sup>[3]</sup>。线上服务更关注“平均停留时间”,是因为在患者线上服务过程中,一位医生可能同时服务多名患者,随着同时服务患者数量越多,导致对单个患者的服务速度变慢,降低每名患者的服务满意度,平均停留时间能对该降低进行更为准确的表征;而“平均等待时间”无法对此进行衡量。本文将各时段到达的线上患者的平均停留时间和线下患者的

平均等待时间的阈值上限分别设定为  $S_M$  和  $W_M$ 。除了服务水平,医院关注其运营成本,本文设定此成本由医生的正常工作时间和加班时间加权构成,这也是所研究问题的优化目标。

## 2 问题模型构建

首先,本文对一个有限周期内的医生排班进行优化,在服务过程假设中根据医生上下班的特点,将一天的工作时间分为了  $|T|$  个时段。医生只在时段开始时上下班,系统需要在每个时段开始时决策每位医生在该时段内是否上班。因此,整个问题是有有限时域、离散时间的。时段  $t$  开始时,系统需要根据当前系统内的医生工作时长和患者等待时间等状态信息  $S_t$ , 决策时段  $t$  内的医生工作  $W_t$ 。时段  $t+1$  开始时的状态只取决于时段  $t$  时的状态和决策,符合马尔可夫性质。因此,以上问题可建模为有限时域内的离散时间马尔可夫决策过程 (discrete-time Markov decision process, DTMDP)。每个时段为 DTMDP 中的每个阶段。在各时段开始时,运营者观察系统中已有的状态并做出决策。DTMDP 的组成要素有状态、动作和状态转移。

### 2.1 系统状态

系统的状态是基于每个时段开始时对系统的观察。在时段  $t$  开始时,状态  $S_t$  定义为  $(H, Q, N, L, M)$ , 其中  $H, Q$  和  $N$  表示各医生的工作状态,  $L$  和  $M$  则表示整个系统内的患者状态,其分别由以下变量组成:

1)  $H_{i,j,t-1}$ : 到时段  $t-1$  末为止,医生  $i$  在诊室  $j$  的连续工作时间。

2)  $Q_{i,t-1}$ : 在时段  $t-1$  中,医生  $i$  所工作的诊室序号(若该医生休息,则序号为0)。

3)  $N_{j,t',t-1}$ : 到时段  $t-1$  末为止,在时段  $t'$  在诊室  $j$  工作的医生数量。

4)  $L_{t-1,t'}$ : 到时段  $t-1$  末为止,在时段  $t'$  到达的线上患者的平均停留时间。

5)  $M_{t-1,t'}$ : 到时段  $t-1$  末为止,在时段  $t'$  到达的线下患者的平均等待时间。

系统的初始状态定义为  $S_0=(0, 0, 0, 0, 0)$ 。

### 2.2 动作集合

各时段的动作决定了医生在该时段的工作状态。设定二元变量  $W_{i,j,t}$  表示医生  $i$  在时段  $t$  内是否在诊室  $j$  工作 ( $j \in J=\{0,1,2\}$ , 其中  $j=1$  和  $j=2$  分别表示在线上 and 线下诊室工作;  $j=0$  表示医生休息)。所选动作需满足医生的排班规则,可行动作空间表示为

$$\sum_{j \in J} W_{i,j,t} = 1, \forall i \in I, t \in T \quad (1)$$

$$\begin{aligned} W_{i,0,t} &= 1, \\ Q_{i,t-1} &= 0, H_{i,1,t-1} > 0, H_{i,2,t-1} > 0, \\ &\forall i \in I, t \in T \end{aligned} \quad (2)$$

$$\begin{aligned} W_{i,0,t} + W_{i,j,t} &= 1, \\ Q_{i,t-1} &= 0, H_{i,j,t-1} = 0, H_{i,3-j,t-1} > 0, \\ &\forall i \in I, j \in \{1,2\}, t \in T \end{aligned} \quad (3)$$

$$\begin{aligned} W_{i,0,t} + W_{i,j,t} &= 1, \\ Q_{i,t-1} &= j, d_j^{\min} \leq H_{i,j,t-1} < d_j^{\max}, \\ &\forall i \in I, j \in \{1,2\}, t \in T \end{aligned} \quad (4)$$

$$\begin{aligned} W_{i,j,t} &= 1, \\ Q_{i,t-1} &= j, H_{i,j,t-1} < d_j^{\min}, \\ &\forall i \in I, j \in \{1,2\}, t \in T \end{aligned} \quad (5)$$

$$\begin{aligned} W_{i,0,t} &= 1, \\ Q_{i,t-1} &= j, H_{i,j,t-1} = d_j^{\max}, \\ &\forall i \in I, j \in \{1,2\}, t \in T \end{aligned} \quad (6)$$

式(1)保证一个医生在时段  $t$  内只有一种工作状态。式(2)~(3)表明,假设医生  $i$  在时段  $t-1$  内休息,如果此医生已经在2个诊室内都工作过,则只能继续休息;如果其只在一个诊室内工作过,那么可以在另一诊室内工作或者继续休息。式(4)~(6)表示,假设医生  $i$  在时段  $t-1$  内在诊室  $j$  工作,若其累计工作时间处于  $[d_j^{\min}, d_j^{\max}]$  内,那么其可以休息或者继续工作;若累计工作时间小于最短时长,则只能继续工作;若工作时间等于最长时长,则只能休息。

### 2.3 状态转移

假设时段  $t$  开始时的系统状态为  $S_t=(H, Q, N, L, M)$ 。系统从2.2节中定义的可行动作空间中选取动作  $W_t=(W_{i,j,t} | i \in I, j \in J)$ , 从状态  $S_t$  转移到状态  $S_{t+1}$ 。向量  $H$  表示了每位医生在各个诊室中所工作的累计时段数。在时段  $t$  中系统通过统计每位医生是否正在工作来更新向量  $H$ 。

$$H_{i,j,0} = 0, \forall i \in I, j \in J \quad (7)$$

$$H_{i,j,t} = W_{i,j,t} \cdot (H_{i,j,t-1} + 1), \forall i \in I, j \in J, t \in T \quad (8)$$

$$d_j^{\min} \leq H_{i,j,t} \leq d_j^{\max}, \quad (9)$$

$$W_{i,j,t} = 0, H_{i,j,t} > 0, \forall i \in I, j \in \{1,2\}, t \in T$$

$$\begin{aligned} d_{\min}^c &\leq H_{i,j,T} \leq d_{\max}^c, \\ H_{i,j,T} &> 0, \forall i \in I, j \in \{1,2\} \end{aligned} \quad (10)$$

式(7)确保在第一个时段开始前,每位医生在各个诊室内工作的时段数为0。式(8)给出了  $H$  的具体更新公式。式(9)和(10)将  $H$  中的元素限定在最短连续工作时长与最长连续工作时长中。向量  $Q$  和  $N$  描述了每位医生在上一个时段的工作状态,它与上个时段动作  $W$  相关。

$$Q_{i,t} = \sum_{j \in J} j W_{i,j,t}, \forall i \in I, t \in T$$

$$N_{j,t',t} = N_{j,t',t-1}, \forall j \in J, t' < t, t \in T$$

$$N_{j,t,t} = \sum_{i \in I} W_{i,j,t}, \forall j \in J, t \in T$$

而对于向量  $L$  和  $M$ , 并没有精确的状态转移公式, 只能通过仿真得到。

$$L_{t,t'} = f_{\text{simu}}(L_{t-1,t'}, W_t) \leq S_M, \forall t' < t, t \in T \quad (11)$$

$$M_{t,t'} = f_{\text{simu}}(M_{t-1,t'}, W_t) \leq W_M, \forall t' < t, t \in T \quad (12)$$

### 2.4 贝尔曼方程

求解马尔可夫决策过程 (Markov decision process, MDP) 中最优决策的重要方法是贝尔曼方程, 它表述了不同阶段状态价值间的递推关系。本文问题目标为最小化医院运营成本, 由正常工作时间和加班时间加权组成, 表示为  $\sum_{i \in I} \sum_{j=1}^2 \sum_{t \in T} W_{i,j,t} + \alpha \sum_{j=1}^2 ot_j$ 。其中, 前项表示各医生正常工作时段总数, 后项为 2 个诊室内医生的加班时间之和,  $\alpha$  为加班时间权重系数。因此 DTM-DP 模型的贝尔曼方程为

$$\begin{cases} V_T^*(S_T) = \min_{W_T \in A_T(S_T)} \left\{ \sum_{i \in I} \sum_{j=1}^2 W_{i,j,T} + \alpha \sum_{j=1}^2 ot_j \right\} \\ V_t^*(S_t) = \min_{W_t \in A_t(S_t)} \left\{ \sum_{i \in I} \sum_{j=1}^2 W_{i,j,t} + V_{t+1}^*(S_{t+1} | S_t, W_t) \right\} \\ 1 \leq t < T \end{cases} \quad (13)$$

在得到以上贝尔曼方程后, 需要据此进行最优排班的求解。动态规划是经典的求解思路, 但考虑到维数爆炸和仿真的复杂性, 其能求解的问题规模极为有限, 需要设计更高效的排班求解和

系统性能评估策略。

## 3 近似动态规划

在动态规划中, 对于阶段  $t$  开始时的状态  $S_t$ , 需要求解贝尔曼最优方程式 (13) 以确定最优的动作  $W_t$ 。在该方程中,  $A_t(S_t)$  表示动作  $W_t$  的可行空间,  $\sum_{i \in I} \sum_{j=1}^2 W_{i,j,t}$  表示系统处于状态  $S_t$  时采取动作  $W_t$  所产生的立即成本,  $S_{t+1}$  为系统处于状态  $S_t$  时采取动作  $W_t$  后转移到的新状态, 而  $V_{t+1}(S_{t+1})$  则是处于状态  $S_{t+1}$  时的状态价值函数。由于存在维数爆炸和仿真的复杂性, 确定最优策略非常耗时, 因此本文提出一种基于 rollout 算法的近似动态规划 (approximate dynamic programming, ADP) 方法对问题进行有效求解<sup>[26]</sup>。rollout 算法采取了前向动态规划的结构, 并在各阶段中通过启发式策略帮助评估每个动作的成本, 从而选取最优的动作<sup>[27-28]</sup>。对于阶段  $t$  中的某个状态  $S_t$ , 定义  $\hat{V}(S_{t+1})$  为通过启发式策略估计的值, rollout 算法选择最优动作的公式为

$$W_t^* = \arg \min_{W_t \in A_t(S_t)} \left\{ \sum_{i \in I} \sum_{j=1}^2 W_{i,j,t} + \hat{V}_{t+1}^*(S_{t+1} | S_t, W_t) \right\}$$

总体算法框架如图 2 所示。在近似动态规划中, 首先根据当前的状态  $S_t$ , 生成所有可行的一步动作  $W_t$ 。在此基础上通过启发式算法搜索每个动作  $W_t$  的后续最优策略, 并将该后续策略的价值作为动作  $W_t$  价值的估算值, 从而为每个阶段选取最优的动作。最后输出最优动作序列作为医生排班问题的最终解。

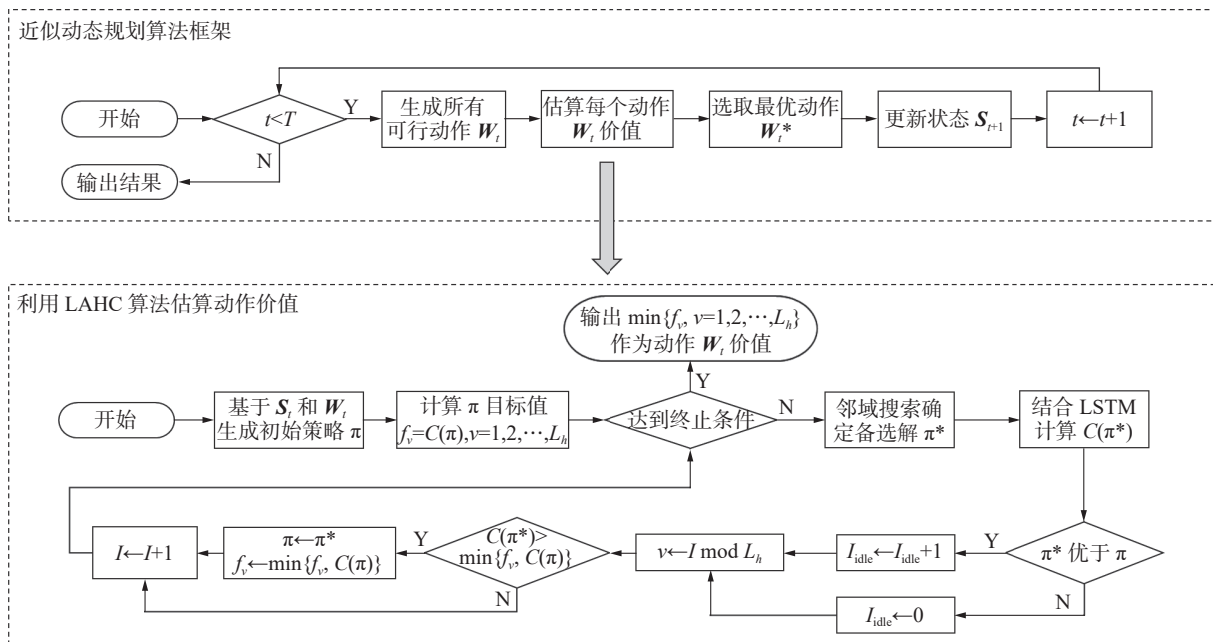


图 2 近似动态规划算法框架

Fig. 2 Framework of approximate dynamic programming algorithm

### 3.1 rollout 算法

根据前视步数的不同, rollout 算法可以分为 one-step-rollout 和 multi-step-rollout。本文根据问题的性质选用了 one-step-rollout 算法, 即当系统到达状态  $S_t$ , 首先对所有可行动作展开评估。对于动作  $W_t$ , rollout 算法调用启发式算法确定选定动作  $W_t$  后的最优策略  $\pi^*(W_t)=(W_{t+1}, W_{t+2}, \dots, W_T)$ 。根据此最优策略计算得到动作  $W_t$  在当前阶段产生的直接价值  $\sum_{i \in I} \sum_{j=1}^2 W_{i,jt}$  和状态  $S_t$  经过动作  $W_t$  转移到新状态  $S_{t+1}$  的价值  $\hat{V}_{t+1}^*(S_{t+1}|S_t, W_t)$ , 从而选取最优动作。因此, 对于状态  $S_t$  一共需要执行  $|A_t(S_t)|$  次启发式算法 ( $|A_t(S_t)|$  为可行动作的数量)。特别地, 当系统在阶段  $t$  末找不到可行动作时, 表明系统在阶段  $t$  所选择的动作是不合适的。为了避免这种情况, 系统将回退到阶段  $t-1$  末状态  $S_{t-1}$ , 重新选择在阶段  $t$  的动作, 前视多步直到阶段  $t+1$ 。在计算得到系统在阶段  $t+1$  存在可行动作后, 系统采用这一重新选取的动作。One-step-rollout 算法的伪代码如下。

输入 初始状态  $S_0$

输出 最优动作序列  $W$

1) 将  $t$  初始化为 0;

2) While  $t < T$  do

3) 根据启发式算法确定状态  $S_t$  下每个可行动作  $W_t$  的价值;

4) if 阶段  $t+1$  时存在可行动作 then;

5)  $W_t^* = \arg \min_{W_t \in A_t(S_t)} \left\{ \sum_{i \in I} \sum_{j=1}^2 W_{i,jt} + \hat{V}_{t+1}^*(S_{t+1}|S_t, W_t) \right\}$

6) 根据  $S_t$  和  $W_t^*$  更新状态  $S_{t+1}$  和  $t$ ;

7) else  $t \leftarrow \max\{0, t-1\}$ ;

8) End while

### 3.2 延迟接受的爬山算法

为了估算每个动作的成本, 本文所提出的 rollout 算法采用了启发式算法来生成可能的后续决策序列, 从而得到每个动作所产生的立即成本与随后的状态价值  $\hat{V}(S_{t+1})$ 。给定状态  $S_t$  和动作  $W_t$ , 通过启发式算法计算后续最优策略  $\pi^*(W_t)=(W_{t+1}, W_{t+2}, \dots, W_T)$ , 该策略给定了时段  $t$  后各医生具体工作状况。结合状态  $S_t$  和最优策略  $\pi^*(W_t)$ , 即得到每个动作  $W_t$  的价值。

本文采用基于延迟接受的爬山算法 (late acceptance hill climbing, LAHC) 作为 rollout 算法中的基础启发式算法<sup>[29]</sup> (即算法 1 中的步骤 5))。与传统爬山法不同, LAHC 算法允许接受非改进的

移动。在每次迭代中, LAHC 算法在确定邻域最优解后, 将其与解池中相应位置 (迭代次数除以解池大小  $L_h$  的余数) 的解进行比较, 从而判断是否移动到该邻域最优解。以  $V^{\pi(W_t)}(S_t)$  表示从状态  $S_t$  出发, 按照策略  $\pi(W_t)$  所确定的医生排班的成本期望值。而 LAHC 算法的目标则在于选定单步前视的动作  $W_t$  后, 求解得到最优的后续策略  $\pi^*(W_t)$ 。LAHC 伪代码如下。

输入 决策前状态  $S$  和所选择的动作  $W$

输出 动作的价值

1) 根据状态  $S$  和  $W$  生成初始策略  $\pi$ ;

2) 计算初始策略的目标值  $C(\pi)$ ;

3) 生成初始解池  $f_v = C(\pi)$ ,  $v = 1, 2, \dots, L_h$ ;

4) 初始化算法参数:  $I \leftarrow 0$ ,  $I_{idle} \leftarrow 0$ ;

5) While  $I \leq I_{maxIters}$  or  $I_{idle} \leq I * 0.02$  do

6) 在  $\pi$  的邻域解中确定邻域最优解  $\pi^*$ ;

7) if  $C(\pi^*) \geq C(\pi)$  then

8)  $I_{idle} \leftarrow I_{idle} + 1$ ;

9) else  $I_{idle} \leftarrow 0$ ;

10)  $v \leftarrow I \bmod L_h$ ;

11) if  $C(\pi^*) < f_v$  or  $C(\pi^*) \leq C(\pi)$  then

12)  $\pi \leftarrow \pi^*$ ;

13) if  $C(\pi) < f_v$  then

14)  $f_v \leftarrow C(\pi)$ ;

15)  $I \leftarrow I + 1$ ;

16) End while

#### 3.2.1 初始解

当处于时段  $t$  时, LAHC 算法需要根据当前的状态  $S_t$  和所选择的动作  $W_t$  生成初始解。初始解生成具体算法如下。1) 检查医生  $i$  是否可以在阶段  $t+1$  以及随后阶段中工作: 如果该医生在阶段  $t+1$  前已经完成了 2 个班次 (一个为线上班次, 一个为线下班次), 则该医生不能够继续工作。否则, 可以在阶段  $t+1$  以及随后阶段为该医生分配班次。2) 如果所选择的动作  $W_t$  中该医生正在某个诊室上班, 则将其班次长度延伸至最大时长  $d_j^{\max}$ ; 如果所选择的动作  $W_t$  中该医生正在休息, 则为其在随机诊室  $j$  内生成一个班次, 具有随机的开始时间  $p$  和结束时间  $q$ 。3) 检查每个诊室在后续时段中是否至少有一个医生值班。如果没有, 对在时段  $t$  内不工作的医生重新随机分配新的班次, 直到满足约束。

#### 3.2.2 邻域结构

在生成初始解后, LAHC 算法将其设置为第一个当前解, 在每次迭代过程中根据当前解生成

邻域解。当前解的所有邻域解通过以下 3 种邻域生成: 1) 如果医生  $i$  没有在诊室  $j$  内的班次, 则额外指派一个在诊室  $j$  内的班次给医生  $i$ ; 2) 如果医生  $i$  有在诊室  $j$  内的班次, 则修改医生  $i$  在诊室  $j$  内的工作时间; 3) 如果医生  $i$  有在诊室  $j$  内的班次, 则取消医生  $i$  在诊室  $j$  内的班次。需要注意的是, 如果邻域解与给定的状态  $S_t$  和动作  $W_t$  冲突, 那么该邻域解将被跳过, 只有符合医生排班约束的邻域解才会被保留。

### 3.2.3 解的评估

在搜索过程中, 当前选择的动作会对服务水平产生长期的影响, 这可能使得式 (11) 和 (12) 难以满足。在 LAHC 算法中对式 (11) 和 (12) 进行松弛, 设定诊室  $j$  在时段  $t$  内的服务水平约束相关的惩罚系数为  $\beta_{j,t}$ , 阶段  $k$  内 LAHC 算法的目标函数为

$$\min \sum_{t=k}^T \left( \sum_{i \in I} \sum_{j \in J \setminus \{0\}} W_{i,j,t} + \beta_{1,t}(S T_t - S_M)^+ + \beta_{2,t}(W T_t - W_M)^+ \right) + \sum_{j=1}^2 o t_j$$

传统的系统性能评估方法往往通过仿真实现。为了保证求解速度, LAHC 算法通过结合数据驱动和仿真方法计算解的目标值。对各邻域解, 首先通过 3.3 节中提出的神经网络预测其系统性能, 若该解的目标值与当前解的目标值差值百分比高于  $\epsilon$ , 则认为该邻域解质量较差, 不再进行精确评估。否则, 通过仿真对该邻域解做精细评估, 从而提高算法搜索精度。

### 3.3 数据驱动的系统性能评估

在 3.2.3 节中, 需要一类方法来对所采用的启发式算法所搜索到的邻域解进行评估, 确定在该邻域解中医生的总工作时间和每个时段到达患者的平均停留/等待时间, 从而判断该邻域解目标值的大小。传统的计算机仿真方法虽然可以实现这个功能, 但需要较长的运行时间。本文设计了深度学习的方法来快速评估给定排班的性能指标。考虑到线上和线下服务系统的差异, 其性能评估通过 2 个相互独立的深度学习模型分别完成。将线上/线下各时段的到达率表示为  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_T\}$ , 各时段在该诊室工作的医生数量表示为  $N = \{N_1, N_2, \dots, N_T\}$ , 二者共同构成了该深度学习模型的特征。对于线上服务系统, 深度学习模型的标签是各时段到达的患者的平均停留时间  $L$  和最

后一个时段的加班时间  $o_{t1}$ , 表示为  $\{L_1, L_2, \dots, L_T, o_{t1}\}$ 。对于线下服务系统, 深度学习模型的标签是各时段到达的患者的平均等待时间  $M$  和最后一个时段的加班时间  $o_{t2}$ , 表示为  $\{M_1, M_2, \dots, M_T, o_{t2}\}$ 。由于输出变量是连续的序列, 该问题是一个序列到序列的回归问题。

#### 3.3.1 模型选择

考虑到问题序列到序列的结构特征, 对系统性能的评估采用长短期记忆 (long short-term memory, LSTM) 的神经网络框架实现<sup>[30]</sup>。LSTM 在处理时间序列型数据方面具有优势, 其利用多个单元间的递归连接, 捕捉和处理来自先前步骤的信息, 并将其用于更新当前的预测或输出, 从而保持数据中序列前后固有的依赖关系<sup>[31-33]</sup>。在单元间, LSTM 相较于传统循环神经网络 (recurrent neural network, RNN) 多了一条承载细胞状态的传输轴, 细胞状态在一个衰减较少的通道里沿时间轴传递, 对时间跨度较大的信息的保持能力比隐藏状态要强很多, 更有效地解决了 RNN 的短期依赖瓶颈。而在单元内, LSTM 通过遗忘门、输入门和输出门对信息进行筛选。LSTM 的特点与本文研究问题结构相吻合, 医生的排班和患者的到达率为时间序列, 同样表现出时间上的依赖性, 前一时刻的系统状态会对后一时刻产生直接影响。LSTM 中输入与输出间的关系具体如下:

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

所使用的神经网络结构如图 3 所示。在 LSTM 模型结构上, 首先通过嵌入层将工作医生人数映射为连续的向量空间中, 提供更好的特征表示。然后在嵌入层基础上, 将工作医生人数与到达率相连接, 共同作为一层 LSTM 层的输入。LSTM 层能够通过隐藏状态和细胞状态捕捉数据中的长短期依赖关系。通过 LSTM 层后, 再经过一个以线性整流函数 (rectified linear unit, ReLU) 为激活函数的全连接层输出最终所需的系统性能指标。全连接层能够对 LSTM 层输出的抽象特征进行进一步的非线性建模, 从而提高模型的表达能力, 使得模型能够更好地拟合复杂的数据分布。

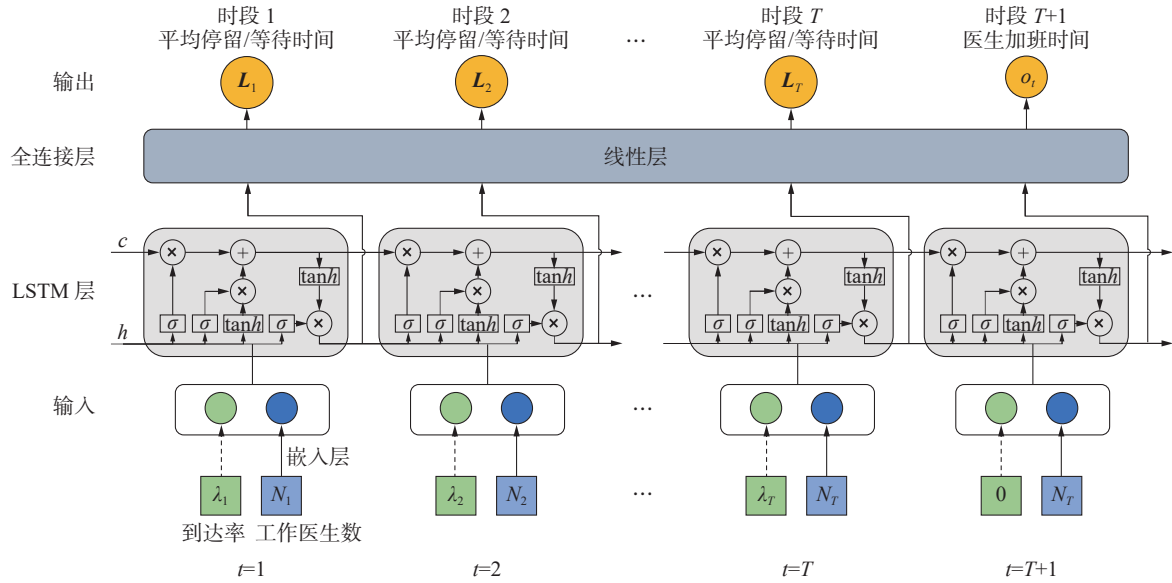


图 3 LSTM 神经网络结构

Fig. 3 Structure of LSTM neural network

3.3.2 数据集准备

神经网络模型的有效性取决于数据集在准确性和广泛适用性方面的质量。通过分析合作医院的实际数据,发现对于一定时期的数据(例如 3 个月内),考虑不同日期,对于一个时段虽然其患者到达率并不相同,但是都基本保持在一定范围内,即不同天的同一时间段内,患者到达率呈现出上下波动的趋势。图 4 为合作医院实际每日患者到达率数据。

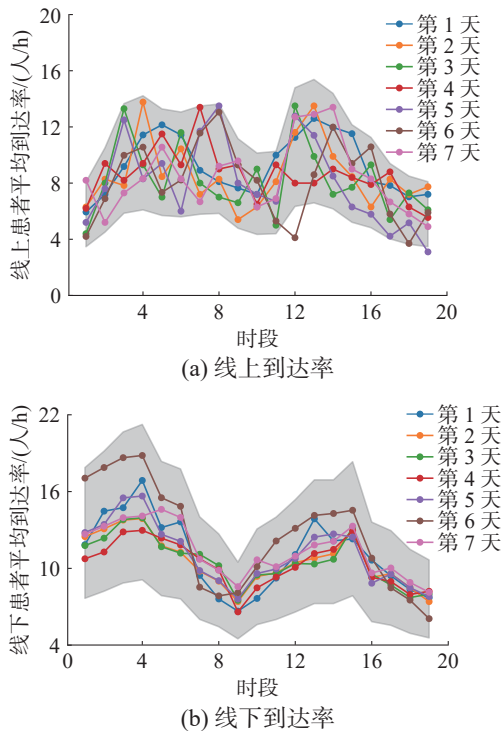


图 4 合作医院实际每日患者到达率数据

Fig. 4 Daily patient arrival rates in cooperative hospital

图 4 给出了合作医院 2023 年 3 月某 7 天的每半个小时到达率数据,从上午 8:00 开始到下午 5:30 共 19 个时段。可见对每个时段而言,其患者到达率处于一定范围内的随机波动。同时,由于可以采集的医院数据量是有限的,为了进行高质量的基于数据驱动方法研究,参考相关研究的方法<sup>[34-35]</sup>,进一步通过计算机仿真生成样本数据以补充数据集。

数据集中每个样本由患者到达率向量  $\lambda$ 、医生人数向量  $N$  和指标向量组成。依照医院实际情况,生成样本步骤如下。首先,如上所述,由于在实际数据中不同日期的对应同时段,患者到达率处于一定范围内波动,因此由实际医院数据得到每个时段的平均到达率,在此基础上将每个时段的到达率乘以一个随机系数,此系数服从  $[0.6, 1.4]$  区间均匀分布,以模拟每天此时段的患者到达率。同时,样本中各时段工作的医生数量也是随机生成,每时段数量在区间  $[1, |I|-1]$  内均匀分布随机产生。需要注意只有在医疗能力充足但不过量的范围内,诊室才能高效运作。因此,所生成的数据要求医生服务能力与患者到达量大致相等,即服务强度必须处于一定的范围内:

$$\rho = \sum_{i=1}^T \lambda_i / \sum_{i=1}^T \mu N_i \in [0.6, 2]$$

如果生成的样本的服务强度不处于此范围内,则重新生成该样本。最后,对每个训练样本,通过蒙特卡洛仿真  $10^5$  次,得到各时段到达患者的平均停留/等待时间,以及最后一个时段的加班时间。所得仿真结果作为每个样本的标签。按照

以上方法重复生成不同样本, 共生成  $10^6$  个样本, 将其按照 4:1 的比例随机划分为训练集与测试集。

### 3.3.3 模型训练

生成数据集后进行神经网络模型训练, 关键任务之一是其超参数设置。考虑到当前问题的规模, 采用了网格搜索的方法。每次迭代都按照特定顺序尝试超参数的组合, 模型根据选定的组合对数据进行拟合, 并记录模型性能, 最后返回性能最优的超参数组合和与之对应的模型。根据网格搜索的结果, 具体所使用超参数如表 1 所示。

表 1 LSTM 所使用超参数  
Table 1 Hyper-parameters of LSTM

超参数	值	含义
嵌入层维度	8×3	医生数映射维数
LSTM层单元数	400	决定模型复杂程度
全连接层单元数	400	决定模型复杂程度
批量大小	8	每次迭代训练时用于更新模型参数的样本数量
学习率	$10^{-4}$	用于控制权重更新步长

本文使用了 AdamW 优化器, 采用小批量梯度下降的算法进行训练。在训练过程中, 通过训练集样本的平均绝对百分比误差 (mean squared error, MSE) 指标来衡量模型性能, 选取最优超参数。假设总共  $n$  条训练数据, 时段长度为  $T+1$  (包含加班时段), 第  $i$  条训练数据在第  $t$  个时段的真实值为  $y_{it}^{true}$ , 预测值为  $y_{it}^{pred}$ , 则 MSE 的计算方式为

$$E_{MSE} = \frac{1}{n(T+1)} \sum_{i=1}^n \sum_{t=1}^{T+1} (y_{it}^{pred} - y_{it}^{true})^2$$

从训练集中随机选取小批量样本用于训练, 最后通过在测试集下的百分比误差判断模型准确程度。

## 4 医院实际数值实验

本节选取合作医院在 2023 年的 7 组实际算例进行数值实验, 数据来自于内科科室, 每个算例对应于一天的数据, 包含实际患者的到达率等信息。该科室中共有 7 名医生, 所使用的日常排班主要构成如下。在线上诊室中, 2 名医生在 8:00~10:30 工作, 另外 2 名医生在 10:30~17:30 工作; 在线下诊室中, 2 名医生在 8:00~14:00 工作, 3 名医生在 14:00~17:30 间工作。基于此 7 个数据算例, 使用本文的方法对医生排班问题进行求解。

首先, 数值实验验证第 3.3 节中所提出系统性能评估方法的精度, 再利用第 3 章 ADP 算法求

解得到新的排班方案, 并对比医院实际排班方案以验证有效性。进一步, 对比 ADP 算法与医生排班问题中常用的模拟退火算法在实际算例中的表现, 以检验 ADP 算法求解的高效性。算法均由 C++ 实现, 运行硬件平台为 3.1 GHz CPU。数值实验中一天医疗服务时间划分为 19 个时段, 每个时段为 0.5 h。单个医生线上最多同时服务 3 个患者, 加班时间成本系数为 2, 线上各时段到达患者平均停留时间阈值为 24 min, 线下各时段到达患者平均等待时间阈值为 45 min。患者到达率、医生服务速率和工作时长限制参数取值均取自医院实际数据和设定。

### 4.1 LSTM 预测精度实验

首先对第 3.3 节所提出的数据驱动的系统性能评估方法进行精度检验, 以检验其在各时段的线上到达患者的平均停留时间、线下到达患者的平均等待时间以及最后一个时段工作医生的加班时间的预测精度。每个算例中包含一天的线上和线下实际到达率, 线上和线下分别根据实际到达率按照 3.3.1 节的方式生成  $2 \times 10^4$  个由随机到达率和随机排班组成的测试集数据。每个测试集数据的特征作为 LSTM 预测模型和离散事件仿真模型的输入, 计算 LSTM 预测模型所输出的系统性能指标的预测值和仿真模型模拟  $10^5$  次后输出的仿真值的平均差值百分比, 结果如表 2 所示。

表 2 LSTM 方法精度 (Y)  
Table 2 Precision(Y) of LSTM neural network %

算例	$s_t$	$ot_1$	$w_t$	$ot_2$
1	0.76	0.76	3.40	1.04
2	0.66	0.76	3.58	1.09
3	0.83	0.87	5.84	1.10
4	0.76	0.89	7.36	1.19
5	0.89	1.19	4.86	0.87
6	1.29	1.40	5.86	1.02
7	0.93	1.10	6.03	1.01
平均	0.87	1.00	5.28	1.05

表 2 中所有列为相应算例的测试集中 LSTM 预测值与仿真结果的平均差值百分比。其中,  $s_t$  列表示线上患者的平均停留时间预测值和仿真值的平均差值百分比,  $ot_1$  列表示线上诊室医生的加班时间预测值和仿真值的平均差值百分比,  $w_t$  列表示线下患者的平均等待时间预测值和仿真值的平均差值百分比,  $ot_2$  列表示线下诊室医生加班时间预测值和仿真值的平均差值百分比。

$$E_{MAPE} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \frac{|y_{i,t}^{pred} - y_{i,t}^{true}|}{y_{i,t}^{true}} \times 100\%$$

可以发现, 对于线上各时段到达患者的平均停留时间和线下各时段到达患者的平均等待时间, LSTM 神经网络预测值与仿真值平均差距分别为 0.87%、5.28%。而对于线上和线下诊室中工作医生的加班时间, 7 个算例中 LSTM 神经网络预测值与仿真值平均差距分别为 1.00% 和 1.05%。这表明所提出的 LSTM 神经网络能够对线上、线下排队系统中的系统性能指标完成高精度的评估。

#### 4.2 求解医生排班的结果

基于医院的 7 个实际算例, 本节采用第 3 章中提出的 ADP 算法求解排班问题, 并把结果和实际排班方案进行对比。ADP 算法中 LAHC 算法迭代次数为 20 次, 解池大小为 5, 服务水平约束相关的惩罚系数为  $10^4$ 。计算结果如表 3 所示。其中  $T_{Break}$  为该排班方案违反阈值约束的次数;  $T_{work}$  为仿真得到的问题目标值, 即医生正常工作时间与加班时间加权之和。第 5 列为实际排班与算法排班目标函数值的百分比相对差距。

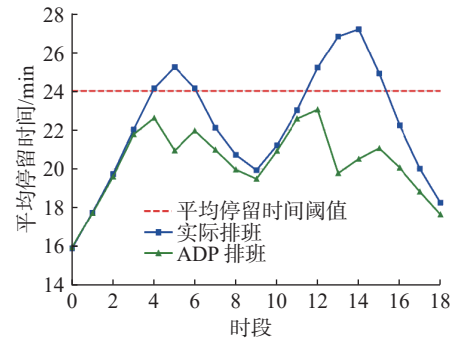
表 3 ADP 所求排班与实际排班比较

Table 3 Comparison of ADP solutions and real-life plan

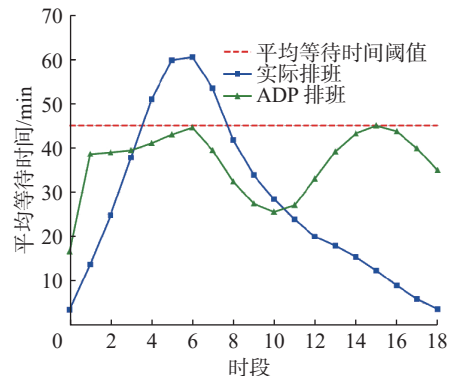
算例	实际排班		算法排班		差距/ %
	$T_{Break}$	$T_{work}$	$T_{Break}$	$T_{work}$	
1	11.0	84.89	0	82.69	2.60
2	6.0	84.80	0	80.35	5.25
3	4.0	84.60	0	76.28	9.83
4	5.0	84.54	0	78.32	7.36
5	6.0	84.21	0	78.04	7.34
6	6.0	84.62	0	80.66	4.68
7	7.0	84.67	0	80.54	4.88
平均	6.4	84.62	0	79.55	5.99

表 3 结果显示, 本文提出的 ADP 算法所得解优于实际排班方案。考虑患者停留/等待时间为代表的服务水平方面, 实际排班平均违反最大停留/等待时间的阈值约束 6.4 次, 而 ADP 算法得到的排班方案中各时段服务水平均满足阈值约束。在目标值方面, ADP 得到的排班方案中的 7 个算例目标值相较于原始排班平均下降了 5.99%, 有效降低了医院的运营成本。可以进一步观察到, 算法所得到的排班方案可以更有效地控制线上患者的平均停留时间和线下患者的平均等待时间。

如图 5 所示, 在算例 1 中, 线上患者的平均停留时间在第 4~6、12~15 时段超出阈值, 线下患者平均等待时间在第 4~7 时段超出阈值。相比之下, 算法排班满足各时段服务水平限制, 并能够在不增加医生容量的情况下提高对患者的服务水平。



(a) 线上平均停留时间



(b) 线下平均停留时间

图 5 算例 1 的 ADP 算法解与实际排班下服务水平对比  
Fig. 5 Comparison of service levels of ADP scheduling solutions and real-life scheduling plan of instance 1

#### 4.3 与其他算法结果对比

本节进一步将所提出的算法与医生排班常用的算法进行比较。模拟退火 (simulated annealing, SA) 算法被证明是解决医生排班问题的有效算法<sup>[8]</sup>。首先将所提出的 ADP 算法和 SA 算法进行排班问题求解的对比。在作为对比的 SA 算法中, 初始解构造、邻域结构、解的评估与 LAHC 算法相同, 但基于 Metropolis 准则判断是否接受新解。初始温度为 100, 降温系数为 0.9, 算法终止条件为 10 次迭代以内当前解未更新则终止。此外, 在 3.3 节中设计了 LSTM 神经网络评估排队系统性能的方法, 替代了传统的仿真方法, 加速 ADP 算法的求解。因此在本节中, 也将所提出的 ADP 算法和传统使用仿真的 ADP 方法进行了比较。在 ADP-仿真算法中, 使用仿真来评估排队系统性能而非 LSTM 神经网络, 其他细节与本文所设计的 ADP 算法保持一致。SA、ADP 和 ADP 仿真方法求解结果对比如表 4 所示, 其中  $T_{work}$  为 3 种算法求解得到的排班方案目标值,  $t$  为 3 种算

法求解所用时间, Gap 计算为  $(T_{work}^{SA} - T_{work}^{ADP})/T_{work}^{ADP}$ 。

表 4 ADP、SA 算法求解结果  
Table 4 Solutions of ADP and SA algorithms

算例	ADP		SA		ADP-仿真		Gap/ %
	$T_{work}$	t/h	$T_{work}$	t/h	$T_{work}$	t/h	
1	82.69	6.8	83.45	10.4	82.69	19.4	0.92
2	80.35	7.4	82.49	11.2	79.40	29.2	2.66
3	76.28	7.1	81.91	8.1	77.05	20.5	7.38
4	78.32	5.6	82.70	10.1	78.32	17.8	5.59
5	78.04	5.6	80.32	6.1	76.92	23.7	2.92
6	80.66	6.2	83.23	7.3	80.01	26.4	3.19
7	80.54	7.5	81.37	8.1	79.44	35.2	1.03
平均	79.55	6.6	82.21	8.8	79.12	24.6	3.40

可以发现, 7 个算例的 ADP 算法解对应的目标函数值相较 SA 算法解平均小 3.40%, 而 ADP 的求解时间相较 SA 算法平均低 128.5 min。7 个算例中 2 种方法的目标函数值最大差距为 7.38%, 最小差距为 0.92%。通过以上对比显示, ADP 算法相较于经典 SA 算法可在合理的时间范围内更有效地降低医院运营成本, 减轻医生工作压力。而在所设计的 ADP 算法和传统 ADP-仿真算法的对比上, 可以发现 ADP 算法相较于 ADP-仿真算法, 目标值的平均差值仅为 0.54%, 求解时间却下降了 73.14%, ADP-仿真算法的求解时间平均达到了 24.1 h。这表明所设计的 ADP 算法能够在更有限的时间内得到高质量的医生排班, 更符合实践应用的需要。

#### 4.4 方法和结果的鲁棒性实验

本文提出的方法中, 一个重要的假设是患者的服务时长服从指数分布。医院实际数据表明, 在部分场景下, 服务时间可能并不适用于指数分布, 更加接近对数正态分布和一般分布(G 分布)。因此, 本节分别设定服务时长服从对数正态分布和 G 分布, 验证所提出的 ADP 算法的鲁棒性。具体方案为按照 ADP 算法生成的排班, 将服务时间设置为对数正态分布和 G 分布, 基于此设定对生成的排班进行系统计算机仿真, 得到系统的性能。对数正态分布和 G 分布的均值和方差均从医院实际数据中得到。仿真所得系统性能如表 5 和 6 所示。从实验结果来看, 当服务时间设定为非指数分布时, ADP 算法所得到的解仍能有效保证服务水平。例如, 当服务时间满足对数正态分布时, 阈值约束违反次数平均从 14.7 次降到了 11.7 次, 医院总运营成本平均降低了 5.98%。

表 5 对数正态分布服务时间下算法解与实际排班结果对比

Table 5 Comparison of ADP solutions and real-life scheduling for lognormal distributed service time

算例	实际排班		算法排班		Gap/ %
	$T_{Break}$	$T_{work}$	$T_{Break}$	$T_{work}$	
1	19.0	85.04	15.0	82.85	2.58
2	14.0	85.01	9.0	80.50	5.31
3	13.0	84.73	11.0	76.36	9.88
4	17.0	84.68	13.0	78.41	7.40
5	13.0	84.25	11.0	78.13	7.26
6	14.0	84.71	12.0	80.82	4.59
7	13.0	84.77	11.0	80.63	4.88
平均	14.7	84.74	11.7	79.67	5.98

表 6 一般分布服务时间下算法解与实际排班结果对比  
Table 6 Comparison of ADP solutions and real-life scheduling for general distributed service time

算例	实际排班		算法排班		Gap/ %
	$T_{Break}$	$T_{work}$	$T_{Break}$	$T_{work}$	
1	14.0	85.06	7.0	83.22	2.17
2	10.0	84.98	5.0	80.84	4.87
3	7.0	84.74	7.0	76.73	9.45
4	8.0	84.69	5.0	78.79	6.97
5	8.0	84.33	4.0	78.47	6.95
6	9.0	84.85	4.0	81.15	4.36
7	9.0	84.89	5.0	81.06	4.51
平均	9.3	84.79	5.3	80.04	5.61

而当服务时间满足 G 分布时, 阈值约束违反次数平均从 9.3 次降到了 5.3 次, 医院总运营成本平均降低 5.61%。这表明, 当服务时间服从其他分布时, 相较于实际排班, ADP 算法所求解排班仍能有效减少医院总运营成本, 控制患者平均停留/等待时长, 具有很强的鲁棒性。

## 5 结束语

在线上 and 线下联合医疗模式下, 面对各自时变的患者到达流, 医院需要通过制定医生的柔性排班方案, 降低运营成本并确保医疗服务水平。为了解决这一问题, 首先针对线上和线下联合医疗服务系统建立了离散时间马尔可夫决策过程模型。在该马尔可夫决策过程模型中, 状态为各时段结束时系统中的患者和医生信息, 动作为各时

段的医生工作指派。为了高效求解马尔可夫决策过程问题,设计近似动态规划算法。接着,提出了基于数据驱动的系统性能评估方法,设计机器学习方法对系统性能高效评估。基于实际数据的数值实验显示,所提出的机器学习方法可快速且高精度完成对系统性能的评估,近似动态规划算法求解得到的排班方案能够有效降低医生工作负荷,并控制患者的停留和等待时间。

本工作可以从多个角度进行扩展。首先,在线下诊室等待队列较长时,一些线下的患者也可能会考虑转移到线上就诊,这将使得问题的建模和求解更具挑战性。同时,本文将每位医生可同时服务的最大患者人数设置为固定值,而在实际服务过程中,可考虑将该值设置为随时间变化的值,以便更容易满足高峰期的患者需求。

## 参考文献:

- [1] 于保荣,杨瑾,宫习飞,等. 中国互联网医疗的发展历程、商业模式及宏观影响因素[J]. 山东大学学报(医学版), 2019, 57(8): 39–52.  
YU Baorong, YANG Jin, GONG Xifei, et al. Development process, business modes and macrocally influencing factors for Internet-based health service in China[J]. Journal of Shandong University (Health Sciences), 2019, 57(8): 39–52.
- [2] 初佃辉,吴军,刘志中,等. 智能化医养融合服务平台关键技术及应用研究[J]. 智能系统学报, 2021, 16(5): 972–988.  
CHU Dianhui, WU Jun, LIU Zhizhong, et al. Research on key technologies and applications of intelligent medical and care integration service platform[J]. CAAI transactions on intelligent systems, 2021, 16(5): 972–988.
- [3] LUO Jun, ZHANG Jiheng. Staffing and control of instant messaging contact centers[J]. Operations research, 2013, 61(2): 328–343.
- [4] RAJAN B, TEZCAN T, SEIDMANN A. Service systems with heterogeneous customers: investigating the effect of telemedicine on chronic care[J]. Management science, 2019, 65(3): 1236–1267.
- [5] 国家互联网信息办公室. 数字中国发展报告(2022年)[EB/OL]. (2023–05–23)[2024–04–02]. [https://www.cac.gov.cn/2023-05/22/c\\_1686402318492248.htm](https://www.cac.gov.cn/2023-05/22/c_1686402318492248.htm).  
Cyberspace Administration of China. Digital China development report(2022)[EB/OL]. (2023–05–23)[2024–04–02]. [https://www.cac.gov.cn/2023-05/22/c\\_1686402318492248.htm](https://www.cac.gov.cn/2023-05/22/c_1686402318492248.htm).
- [6] ZHOU Cuihua, HAO Yifei, LAN Yanfei, et al. To introduce or not? strategic analysis of hospital operations with telemedicine[J]. European journal of operational research, 2023, 304(1): 292–307.
- [7] 罗利,付颖,徐雪茹,等. 多方参与的互联网医院运营管理研究综述[J]. 工业工程与管理, 2024, 29(4): 205–222.  
LUO Li, FU Ying, XU Xueru, et al. The operation management of multi-participant internet hospital: a framework and review[J]. Industrial engineering and management, 2024, 29(4): 205–222.
- [8] ERHARD M, SCHOENFELDER J, FÜGENER A, et al. State of the art in physician scheduling[J]. European journal of operational research, 2018, 265(1): 1–18.
- [9] YOUN S, GEISMAR H N, PINEDO M. Planning and scheduling in healthcare for better care coordination: current understanding, trending topics, and future opportunities[J]. Production and operations management, 2022, 31(12): 4407–4423.
- [10] BRUNNER J O, BARD J F, KOLISCH R. Midterm scheduling of physicians with flexible shifts using branch and price[J]. IIE transactions, 2010, 43(2): 84–109.
- [11] STOLLETZ R, BRUNNER J O. Fair optimization of fortnightly physician schedules with flexible shifts[J]. European journal of operational research, 2012, 219(3): 622–629.
- [12] WANG Zixiang, LIU Ran, SUN Zhankun. Physician scheduling for emergency departments under time-varying demand and patient return[J]. IEEE transactions on automation science and engineering, 2023, 20(1): 553–570.
- [13] SINREICH D, JABALI O, DELLAERT N P. Reducing emergency department waiting times by adjusting work shifts considering patient visits to multiple care providers [J]. IIE transactions, 2012, 44(3): 163–180.
- [14] LIU Ran, XIE Xiaolan. Physician staffing for emergency departments with time-varying demand[J]. INFORMS journal on computing, 2018, 30(3): 588–607.
- [15] ZAERPOUR F, BIJVANK M, OUYANG Huiyin, et al. Scheduling of physicians with time-varying productivity levels in emergency departments[J]. Production and operations management, 2022, 31(2): 645–667.
- [16] 祖漫漫. 门诊医生排班问题研究[D]. 大连: 大连理工大学, 2019.  
ZU Manman. Study on outpatient doctors' scheduling problem[D]. Dalian: Dalian University of Technology, 2019.
- [17] 王铖恺,范晓宇,徐捷,等. 结合 Benders 分解和列生成的发热门诊排班数学建模和优化算法[J]. 系统管理学报, 2023, 32(3): 476–487.  
WANG Chengkai, FAN Xiaoyu, XU Jie, et al. Mathematical modeling and optimization algorithm for fever clinics scheduling combining benders decomposition and column generation[J]. Journal of systems & management, 2023, 32(3): 476–487.
- [18] 兰绍雯. 门诊医生资源协同优化问题研究[D]. 合肥: 合肥工业大学, 2022.  
LAN Shaowen. Research on collaborative optimization of

- outpatient doctor resources[D]. Hefei: Hefei University of Technology, 2022.
- [19] PRAKASH A M, ZHONG Xiang. Queueing information management of primary care delivery with E-visits[J]. *IEEE transactions on automation science and engineering*, 2022, 19(2): 632–645.
- [20] ÇAKIÇI Ö E, MILLS A F. On the role of teletriage in healthcare demand management[J]. *Manufacturing & service operations management*, 2021, 23(6): 1483–1504.
- [21] 姜劲, 白闪闪, 王云婷, 等. 线上和线下医疗服务质量对患者线下就医决策的影响[J]. *管理科学*, 2020, 33(1): 46–53.
- JIANG Jin, BAI Shanshan, WANG Yunting, et al. Influence of online and offline medical service quality on patients' offline medical decision-making[J]. *Journal of management science*, 2020, 33(1): 46–53.
- [22] HOSSAIN I S, AKHAND M A H, SHUVO M I R, et al. Optimization of university course scheduling problem using particle swarm optimization with selective search[J]. *Expert systems with applications*, 2019, 127: 9–24.
- [23] 宋婷, 陈矛, 吴超, 等. 基于多类迭代局部搜索的自动化排课算法[J]. *计算机应用*, 2019, 39(6): 1760–1765.
- SONG Ting, CHEN Mao, WU Chao, et al. Automated course arrangement algorithm based on multi-class iterated local search[J]. *Journal of computer applications*, 2019, 39(6): 1760–1765.
- [24] GUPTA V, ZHANG Jiheng. Approximations and optimal control for state-dependent limited processor sharing queues[J]. *Stochastic systems*, 2022, 12(2): 205–225.
- [25] CUI L, TEZCAN T. Approximations for chat service systems using many-server diffusion limits[J]. *Mathematics of operations research*, 2016, 41(3): 775–807.
- [26] POWELL W B. Approximate dynamic programming: Solving the curses of dimensionality[M]. New York: John Wiley & Sons, 2007.
- [27] POWELL W B. Perspectives of approximate dynamic programming[J]. *Annals of operations research*, 2016, 241(1): 319–356.
- [28] LIU Derong, XUE Shan, ZHAO Bo, et al. Adaptive dynamic programming for control: a survey and recent advances[J]. *IEEE transactions on systems, man, and cybernetics: systems*, 2021, 51(1): 142–160.
- [29] BURKE E K, BYKOV Y. The late acceptance hill-climbing heuristic[J]. *European journal of operational research*, 2017, 258(1): 70–78.
- [30] YU Yong, SI Xiaosheng, HU Changhua, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. *Neural computation*, 2019, 31(7): 1235–1270.
- [31] 张驰, 郭媛, 黎明. 人工神经网络模型发展及应用综述[J]. *计算机工程与应用*, 2021, 57(11): 57–69.
- ZHANG Chi, GUO Yuan, LI Ming. Review of development and application of artificial neural network models[J]. *Computer engineering and applications*, 2021, 57(11): 57–69.
- [32] 李倩玉, 王蓓, 金晶, 等. 基于双向 LSTM 卷积网络与注意力机制的自动睡眠分期模型[J]. *智能系统学报*, 2022, 17(3): 523–530.
- LI Qianyu, WANG Bei, JIN Jing, et al. Automatic sleep staging model based on the bi-directional LSTM convolutional network and attention mechanism[J]. *CAAI transactions on intelligent systems*, 2022, 17(3): 523–530.
- [33] 常明煜, 田乐, 郭茂祖. 基于 HHT-LSTM 的冬奥会临时设施运行趋势预测方法研究[J]. *智能系统学报*, 2024, 19(1): 228–237.
- CHANG Mingyu, TIAN Le, GUO Maozu. Research on the HHT-LSTM-based operation trend prediction method of temporary facilities for the Winter Olympic Games[J]. *CAAI transactions on intelligent systems*, 2024, 19(1): 228–237.
- [34] XAVIER Á S, QIU Feng, AHMED S. Learning to solve large-scale security-constrained unit commitment problems[J]. *INFORMS journal on computing*, 2021, 33(2): 739–756.
- [35] WANG Zixiang, LIU Ran. Managing appointments of outpatients considering the presence of emergency patients: the combination of the analytical and data-driven approach[J]. *International journal of production research*, 2022, 60(13): 4214–4228.

#### 作者简介:



张越, 硕士研究生, 主要研究方向为基于机器学习与人工智能的服务系统优化。E-mail: caesar8310@sjtu.edu.cn。



王子翔, 讲师, 博士, 主要研究方向为服务系统建模优化和基于数据驱动的医疗系统运作管理。E-mail: zixiang\_wang@hznu.edu.cn。



刘冉, 副教授, 博士生导师, 主要研究方向为生产与服务系统运作管理, 发表学术论文 60 余篇。E-mail: liuran2009@sjtu.edu.cn。