



## 基于预测与强化学习的5G混合切片资源优化分配

陈智雄, 谢宇鹏, 郭以贺

引用本文:

陈智雄, 谢宇鹏, 郭以贺. 基于预测与强化学习的5G混合切片资源优化分配[J]. *智能系统学报*, 2026, 21(3): 739–750.

CHEN Zhixiong, XIE Yupeng, GUO Yihe. Prediction and reinforcement learning-based optimized resource allocation for 5G hybrid network slicing[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(3): 739–750.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202506012>

## 您可能感兴趣的其他文章

### 一种卷积神经网络集成的多样性度量方法

Diversity measuring method of a convolutional neural network ensemble

智能系统学报. 2021, 16(6): 1030–1038 <https://dx.doi.org/10.11992/tis.202011023>

### 多元商品智能交易系统

Multi-commodity intelligent transaction system

智能系统学报. 2020, 15(6): 1188–1196 <https://dx.doi.org/10.11992/tis.201909007>

### 加权PageRank改进地标表示的自编码谱聚类算法

An autoencoder spectral clustering algorithm for improving landmark representation by weighted PageRank

智能系统学报. 2020, 15(2): 302–309 <https://dx.doi.org/10.11992/tis.201904021>

### 弹性网络核极限学习机的多标记学习算法

Multi-label learning algorithm of an elastic net kernel extreme learning machine

智能系统学报. 2019, 14(4): 831–842 <https://dx.doi.org/10.11992/tis.201806005>

### 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>

### 基于深度学习的视频预测研究综述

Review of deep learning-based video prediction

智能系统学报. 2018, 13(1): 85–96 <https://dx.doi.org/10.11992/tis.201707032>

DOI: 10.11992/tis.202506012

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20260304.1803.002>

# 基于预测与强化学习的 5G 混合切片资源优化分配

陈智雄<sup>1,2,3</sup>, 谢宇鹏<sup>1</sup>, 郭以贺<sup>1</sup>

(1. 华北电力大学 电子与通信工程系, 河北 保定 071003; 2. 华北电力大学 河北省电力物联网技术重点实验室, 河北 保定 071003; 3. 华北电力大学 电力物联智能化技术河北省工程研究中心, 河北 保定 071003)

**摘要:** 针对现有算法在切片资源合理分配和服务水平协议 (service-level agreement, SLA) 保障中的学习频次过高、计算复杂度较大等问题, 提出了一种融合时间序列预测与强化学习的智能时频资源分配策略。考虑 SLA 违反概率约束和 5G 时频资源等约束, 以时频资源使用数量最小化为目标, 建立了 5G 混合场景的切片资源优化分配模型。首先利用堆叠长短期记忆网络 (stacked long short-term memory, Stacked-LSTM) 对 eMBB、URLLC、mMTC 场景的信噪比、排队缓冲区容量、时延、设备连接数量的 SLA 关键指标进行预测; 定义了高斯核强化学习的状态、动作和奖励函数, 并利用 SLA 指标预测值和随机启动概率门限执行学习算法获得准最佳分配策略。相比于现有算法, 仿真结果表明所提算法可以在保障 SLA 违背概率约束和时频资源数量等基本性能的前提下, 在预测精度、收敛速度和资源分配等综合性能上表现较好, 有效地减少机器学习的频次和复杂度。

**关键词:** 5G 混合场景; 切片; 空口资源分配; 服务水平协议; 强化学习; 预测; 随机启动概率门限; 学习频次

中图分类号: TP393; TN929.5 文献标志码: A 文章编号: 1673-4785(2026)03-0739-12

中文引用格式: 陈智雄, 谢宇鹏, 郭以贺. 基于预测与强化学习的 5G 混合切片资源优化分配 [J]. 智能系统学报, 2026, 21(3): 739-750.

英文引用格式: CHEN Zhixiong, XIE Yupeng, GUO Yihe. Prediction and reinforcement learning-based optimized resource allocation for 5G hybrid network slicing[J]. CAAI transactions on intelligent systems, 2026, 21(3): 739-750.

## Prediction and reinforcement learning-based optimized resource allocation for 5G hybrid network slicing

CHEN Zhixiong<sup>1,2,3</sup>, XIE Yupeng<sup>1</sup>, GUO Yihe<sup>1</sup>

(1. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071003, China; 2. Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, China; 3. Hebei Engineering Research Center of Intelligent Technology for Power Internet of Things, North China Electric Power University, Baoding 071003, China)

**Abstract:** To address the high learning frequency and high computational complexity of existing algorithms for rational slice resource allocation and service-level agreement (SLA) assurance, this paper proposes an intelligent time-frequency resource allocation strategy that integrates time series prediction with reinforcement learning. Under constraints on SLA violation probability and 5G resources, an optimization model for slice resource allocation in hybrid 5G scenarios is established, with the objective of minimizing the time-frequency resource usage. First, a stacked long short-term memory network is employed to predict key SLA indicators such as signal-to-noise ratio, queue buffer occupancy, latency, and the number of connected devices for enhanced mobile broadband, ultra-reliable low-latency communications, and massive machine-type communications scenarios. Next, the states, actions, and reward function of the Gaussian-kernel-based reinforcement learning algorithm are defined. The learning process is guided by the predicted SLA indicators and a randomly initialized probabilistic gating threshold to derive a near-optimal resource allocation policy. Simulation results demonstrate that, compared with existing algorithms, the proposed approach achieves better overall performance in prediction accuracy, convergence speed, and resource allocation efficiency while ensuring basic performance metrics such as SLA violation probability and resource usage. It also reduces the frequency and complexity of machine learning operations.

**Keywords:** 5G hybrid scenarios; network slice; radio access resource allocation; service-level agreement; reinforcement learning; predictive; randomized initialization probabilistic threshold; learning frequency

收稿日期: 2025-06-12. 网络出版日期: 2026-03-05.  
基金项目: 国家自然科学基金青年基金项目 (61601182).  
通信作者: 陈智雄. E-mail: [zxchen@ncepu.edu.cn](mailto:zxchen@ncepu.edu.cn).

随着工业互联网的加速发展, 众多车间设备日益接入工厂内部网络。尤其是自动引导小车 (automated guided vehicle, AGV) 等移动终端对通

信系统在时延、灵活性与带宽方面提出了更高要求,传统有线网络已难以满足其对广泛连接和高移动性的需求<sup>[1-2]</sup>。5G 凭借其其在混合场景方面的优势,为突破工业物联网发展瓶颈提供了关键思路。同时,智能制造快速推进使设备类型多样化、接入量剧增,5G 网络面临高度定制化且复杂的服务需求<sup>[3-4]</sup>。然而传统网络架构由于耦合度高、灵活性不足,难以实现对多样化业务流的实时调度与资源的精细化管理,已无法有效适应工业场景中频繁变化的通信需求。在此背景下,5G 切片技术应运而生<sup>[5-6]</sup>,作为支撑多样化业务、高效配置资源的核心技术,成为运营商和设备厂商关注的焦点。

资源分配与管理是 5G 网络切片领域的核心研究问题之一<sup>[7]</sup>。然而在实际应用中,智能制造场景的高度动态性和复杂性,尤其是在混合业务场景下多任务并行执行的背景下,传统静态资源分配策略已难以满足其性能需求<sup>[8]</sup>。为实现多业务在共享网络环境下的资源最优配置,提升网络资源利用率并有效控制运营成本,动态资源分配策略逐渐成为该领域研究的重点方向<sup>[9-10]</sup>。

近年来学者们开始探索基于机器学习的 5G 网络切片资源分配方法。在 5G 通信体系中,文献 [11] 提出一种基于优先级的机制,实现了在资源竞争的切片之间的性能隔离。文献 [12] 提出了联邦元强化学习框架,在总体成本、端到端时延及收敛速度等方面表现出良好性能。文献 [13-14] 提出将深度学习与确定性网络切片实例策略融合的优化算法,提升了切片资源分配方案的可扩展性、灵活性和鲁棒性,并减少切片实例的平均队列等待时间。文献 [15] 则提出了一种基于深度迁移强化学习的策略,用于 5G 无线接入网 (radio access network, RAN) 切片中联合无线和缓存资源分配,实现较低的时延和较高的系统吞吐量。文献 [16] 中提出一种基于机器学习的 RAN 切片分配策略,在保证各种切片业务的服务满意度同时,最大限度地提高频谱使用效率。

尽管上述研究在不同程度上提升了切片资源分配的性能,但均依赖于无模型强化学习方法。此类方法虽具备直观策略设定的简便性、实现与训练相对容易等优势,但作为纯粹基于经验驱动的优化框架,仍面临以下挑战:算法可解释性差、样本效率低、策略空间探索冗余以及难以处理多目标冲突等问题<sup>[17]</sup>。尤其在面向检测流水线等动态变化显著、数据复杂度高的智能制造环境中,若仍采用无模型强化学习方法,可能导致资源使

用效率低下并频繁违反 SLA 约束,难以满足关键检测终端的业务需求。

相较于无模型方法,有模型强化学习通过利用动态模型生成并利用经验数据进行学习,展现出更强的泛化与适应能力。文献 [18] 提出了一种基于模型的强化学习算法,在 SLA 保障、资源利用效率以及计算成本控制方面,优于部分无模型强化学习方法。文献 [19] 则深入分析队列配置对流量限制和丢包率等指标的影响,并利用基于模型的强化学习算法生成优先级、公平性与拥塞控制的优化策略。文献 [20] 提出了一种基于模型的强化学习框架,通过引入奖励替代模型显著减少了频繁的性能评估的计算成本。文献 [21] 提出了一种基于模型可信度的强化学习方法,有效提升了样本的利用效率以及避免了模型偏差的累积。

然而在检测流水线中,资源分配面临的不仅是多终端之间的合理协调问题。随着业务规模的不断扩大,资源配置需结合业务数据的动态特征实现更为灵活的调度<sup>[22]</sup>。现有的算法在面对高频、复杂任务下快速且精准地完成资源分配方面仍存在显著不足。

为了更加高效、可扩展性地解决资源分配问题,资源预测分配提供了一种新的解决思路。文献 [23-24] 提出了面向多切片管理的动态阈值算法来预测切片需求及拥塞概率,动态选择合适的控制策略集,以此提升整体管理效率。文献 [25] 提出一种基于预测的边缘资源垂直缩放算法,能够减少危险预警服务的服务水平目标违规。文献 [26] 则提出将强化学习和长短期记忆网络相结合的综合预测方法,在数据稀疏的情况下取得高精度的多元负荷预测结果。文献 [27-28] 在空间和时间维度上提出了一种自动特征选择的预测框架,可在大量数据下提高预测方案的准确性。

尽管上述研究在切片需求预测与网络资源优化方面取得了积极进展,但仍存在诸多限制。例如多数预测算法复杂度较高,缺乏对智能体学习频率与训练开销的系统性考量;同时在算力受限的部署环境下,易导致系统学习周期过长,影响其在实际场景中的可部署性与实时响应能力<sup>[29]</sup>。鉴于在 5G 混合场景下实现切片资源合理分配和满足 SLA 保障的需求,针对现有算法普遍面临学习频次过高、计算复杂度过大等问题,本文面向载波通信模块检测流水线网络,考虑 eMBB、URLLC 和 mMTC 混合业务场景中服务水平协议违反概率和 5G 空口时频资源总量有限等约束,建立了以时频资源使用数量最小化为目标的 5G

混合切片资源优化分配模型。为解决优化分配频次和复杂度与性能的平衡问题, 提出了一种基于 Stacked-LSTM 预测与高斯核强化学习的混合切片资源分配算法, 结合状态预测值和随机启动概率门限来执行学习算法。进一步, 本文定义了高斯核强化学习的状态、动作和奖励函数, 并给出了高斯核强化学习训练和决策的关键步骤, 利用仿真实验验证了预测算法的精度, 以及混合切片资源优化分配算法的性能。

## 1 系统模型

### 1.1 载波通信模块检测流水线的 5G 专网设计

载波通信模块检测流水线中存在一些检测终端和转运设备, 如 AGV、信息采集设备、性能检测单机等, 由于传统 4G 网络无法满足这些终端所需的大带宽、海量设备接入、低时延数据传输等条件, 故本文基于 5G 网络架构对载波通信模块检测流水线进行 5G 专网设计。具体网络架构如图 1 所示。

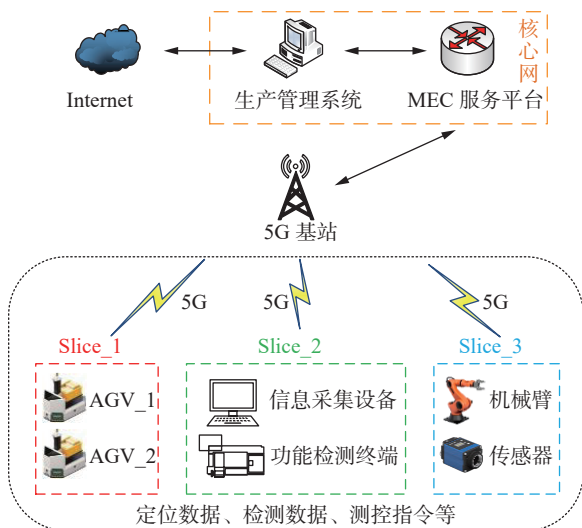


图 1 通信模块自动化检测流水线 5G 网络架构

Fig.1 5G network architecture for automated testing pipelines of communication modules

在该流水线 5G 网络架构中, 接入网部分由 5G 工业物联网关 CPE(customer premises equipment)、5G 基站构成, 核心网部分由 MEC(multi-access edge computing) 边缘服务平台和生产管理系统构成。其中 5G 工业物联网关 CPE 主要负责连接 5G 业务终端, 是 5G 通信中较为关键的环节。关于接入 5G 专网的终端设备通信流程, 以 AGV 为例: AGV 在执行运输任务时, 实时将自身的定位数据向 5G 工业物联网关 CPE 发送; 5G 工业物联网关 CPE 再通过 5G 基站向 MEC 边缘服务平台转发定位数据; 最后由生产管理系统统一安排

任务, 进行调度优化。

### 1.2 5G 网络切片系统模型

由于多个终端接入 5G 专网时容易造成对时频资源的抢夺占用, 为了实现 5G 各种不同业务之间的资源相互隔离与数据传输安全, 考虑采用 5G 网络切片技术。

考虑一个流水线厂房的多切片下行传输场景, 该场景由一个 5G 基站、多个终端组成, 图 2 为系统网络示意图。不同业务类型的终端具有不同的数据包大小、包到达时间间隔, 以及不同的 SLA 需求。基站包含的  $B$  个切片共享该基站能够提供的所有无线资源。将基站和终端之间的无线电接口处构造为帧, 每帧划分为时间和频率维度: 在时间维度上, 帧划分为传输时间间隔 (transmission time intervals, TTIs), 假设 TTI 持续时间为 1 ms; 在频率维度上, 帧被划分为频率子载波, 假设子载波间隔为 15 kHz。其中最小的时频组合单元称为资源块 (resource block, RB), 由 1 个 TTI 和 12 个频率子载波组成, 这些基本的资源块共同构成了整个无线频谱资源。本文考虑给每个切片分配对应的时频资源, 并且这些资源由该切片单独使用, 从而保证了切片之间资源的相互隔离。设  $B$  为射频接口中共存的  $B$  个网络切片集合, 在  $B$  个切片的系统场景中,  $C$  为在这些切片之间分配的资源块总数, SLA 违反数量以  $\delta$  为上界。因此本文所提的 5G 混合切片资源优化分配模型的最优资源分配策略目标为

$$\begin{aligned} & \min \sum_{i=1}^B RBs_{(i)}, \text{ for } i \in \mathcal{B} \\ & \text{s.t. } \sum_{i \in \mathcal{B}} SLA_{\text{violation}} \leq \delta, \sum_{i \in \mathcal{B}} RBs_{(i)} \leq C \end{aligned} \quad (1)$$

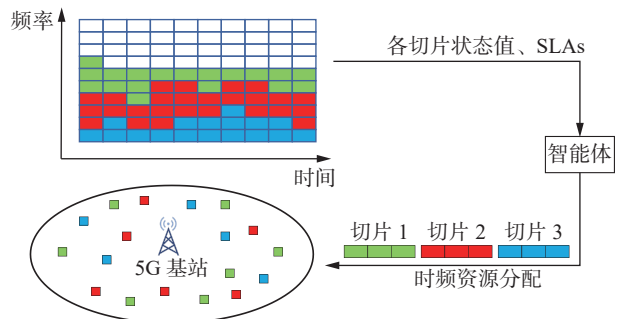


图 2 系统网络

Fig.2 System network

## 2 Stacked-LSTM-GKBRL 算法

在 RAN 切片的时频资源分配问题中, 为了满足各切片的 SLA, 分配过多的时频资源容易造成时频资源的浪费, 导致无法合理利用时频资源;

而且时频资源总数有限, 分配过少的时频资源又无法保障满足各切片的 SLA。故针对此问题, 为了使分配资源的数量最小化, 合理分配资源的同时还需要尽量减少每个切片的 SLA 违反数量, 因此本文提出 Stacked-LSTM-GKBRL 算法。

### 2.1 Stacked-LSTM-GKBRL 模块

为了避免在拥塞高峰期时段中众多检测终端争抢资源的情况发生, 保证不同设备的可靠连接和及时的数据传输, 本文提出将堆叠长短期记忆

网络预测算法与高斯核强化学习 (Gaussian-kernel-based reinforcement learning, GKBRL) 算法融合的混合切片资源算法来探索切片各 SLA 指标变化规律。两种算法的结合能够有效辅助智能体对网络环境进行感知, 提前预测并适应各 SLA 指标的变化, 从而实现更加智能的资源调度和系统性能优化。将 Stacked-LSTM 预测算法和 GKBRL 算法相结合, 构建资源分配模型, 简称为 Stacked-LSTM-GKBRL 模型, 算法架构如图 3 所示。

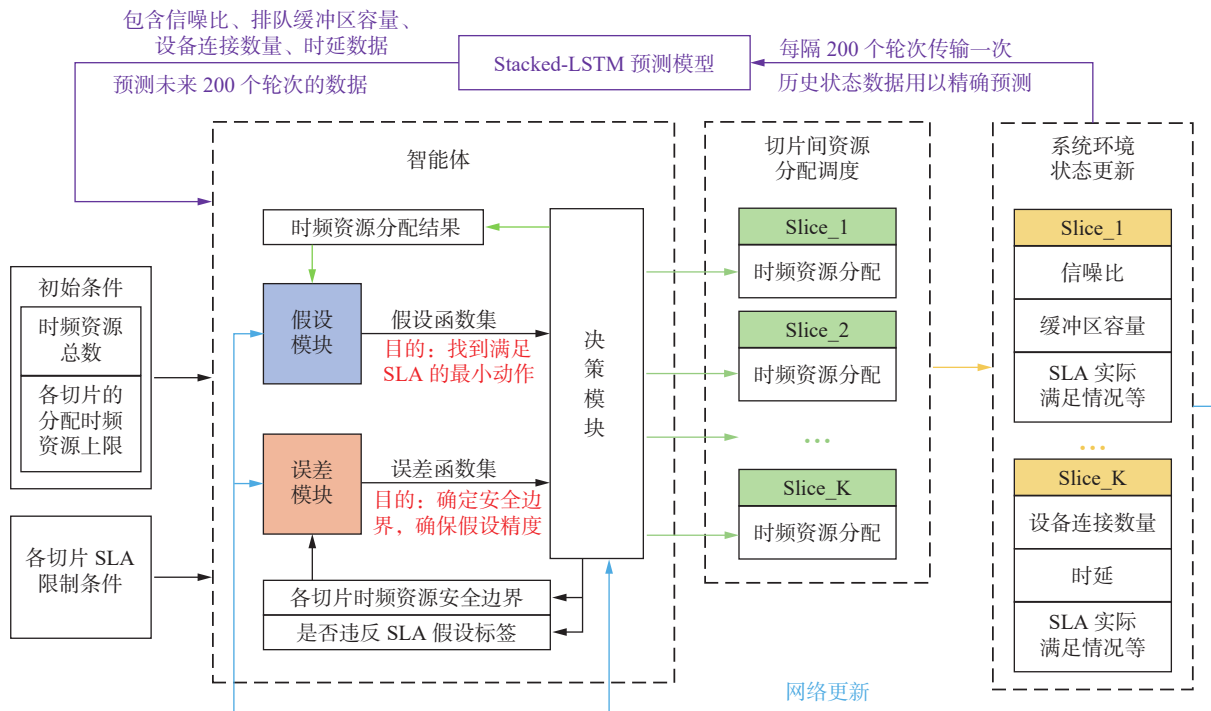


图 3 Stacked-LSTM-KBRL 算法示意

Fig. 3 Schematic of the Stacked-LSTM-KBRL algorithm

如图 3 所示, Stacked-LSTM-GKBRL 模型具体由 Stacked-LSTM 预测模块、假设模块、误差模块和决策模块共 4 个模块构成。首先初始化数据, 智能体先与系统环境交互 200 轮次后, 将记录下的历史状态数据输送到 Stacked-LSTM 预测模块中。Stacked-LSTM 预测模块对切片的各项 SLA 指标, 包含信噪比、排队缓冲区容量、时延、连接设备数量 4 项指标, 对未来 200 轮次进行预测, 并将预测数据传输至智能体中做进一步处理。其中 Stacked-LSTM 预测模块每 200 轮次执行一次。

### 2.2 Stacked-LSTM 预测模块

由于传统长短期记忆网络 (long short-term memory, LSTM) 模型参数相对较少, 导致对数据变化的敏感性高, 鲁棒性和泛化能力差<sup>[30]</sup>。而相较于其他预测算法, Stacked-LSTM 算法能够有效捕捉复杂序列数据之间的长期依赖性, 从而克服

在处理长期依赖和梯度消失问题上的困难。考虑复杂度、计算时间等成本因素, Stacked-LSTM 算法不会与 GKBRL 算法共同执行每个轮次的运算, 而只是预测出一定数量轮次的序列结果, GK-BRL 算法再根据序列结果进行轮次学习和后续时频资源分配。

Stacked-LSTM 包括 3 个门: 遗忘门、输入门和输出门, 图 4 为 Stacked-LSTM 结构中的一层。其中每个门还包含一个单独的记忆细胞, 用来记录长期信息。遗忘门指示从记忆细胞中丢弃一些信息。输入门负责将一些新信息存储在记忆细胞中。第 3 个是输出门, 在传输时间间隔为 LSTM 模块提供最终输出。每一轮次的输出结果将作为下一轮次的输入, 继续预测迭代, 直至 200 轮次循环结束。最后将结果处理为状态序列, 作为智能体的输入。

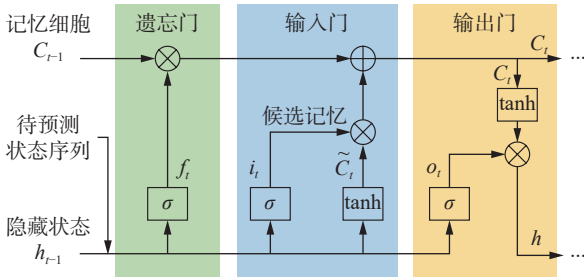


图 4 Stacked-LSTM 预测算法结构

Fig. 4 Architecture of the Stacked-LSTM prediction algorithm

Stacked-LSTM 算法还使用均方误差  $E_{MS}$  (mean squared error, MSE) 作为损失函数, 其衡量了预测值与真实值之间的差异, 计算的是预测误差的平方的平均值:

$$E_{MS} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

式中:  $n$  为样本总数,  $y_i$  为第  $i$  个样本的真实值,  $\hat{y}_i$  为模型预测的第  $i$  个样本的值。同时, 本文中 will 使用自适应矩估计 (adaptive moment estimation, Adam) 作为优化器, 通过不断优化损失函数, 帮助模型尽可能地输出准确的结果。

### 2.3 预测数据处理

智能体接收 Stacked-LSTM 预测模块传来的预测数据后, 首先对各项指标数据进行扫描检测。若存在情况: 选择相同的轮次作为起始点, 之后的  $Y$  ( $Y \leq 25$ ) 个连续轮次的数据都满足 SLA; 智能体则基于这  $Y$  个连续轮次的数据组成该区间的状态量, 并记录这个区间的起始轮次  $n_{start}$  和结束轮次  $n_{end}$ 。其次, 针对  $Y$  个连续轮次的数据都满足 SLA 的情况, 将预测数据分为 4 组序列: 信噪比  $\{1, 2, \dots, Y_s\}$ 、排队缓冲区容量  $\{1, 2, \dots, Y_q\}$ 、连接设备数量  $\{1, 2, \dots, Y_c\}$  和时延  $\{1, 2, \dots, Y_d\}$ , 每组序列的  $Y$  个数据都满足 SLA, 并从各组序列中筛选出相对性能最差的量: 信噪比  $y_s$ 、排队缓冲区容量  $y_q$ 、时延  $y_d$  和连接设备数量  $y_c$ 。随后, 将这 4 个相对性能最差的量组成新的状态  $s \leftarrow \{y_s, y_q, y_d, y_c\}$ 。具体流程如图 5 所示。

最后, 智能体基于新状态  $s$  执行 GKBRL 算法获得动作  $a$ , 用作这  $Y$  个连续轮次的资源分配。其中, 在这  $Y$  个连续轮次内智能体将不再进行强化学习, 而是每个轮次都执行动作  $a$ , 以及记录执行动作  $a$  后的系统环境状态, 为后续 Stacked-LSTM 模型预测提供数据。同时, 算法还考虑到随机性因素, 为了增加 Stacked-LSTM-GKBRL 模型的随机性还设置了随机启动概率门限  $p_{start}$ 。在每个轮次执行动作  $a$  后抽取一个随机

数, 若小于随机启动概率门限, 则智能体停止在下一轮次执行动作  $a$ , 并基于之前执行动作  $a$  后观察到的系统环境状态启动强化学习。具体如图 6 所示。

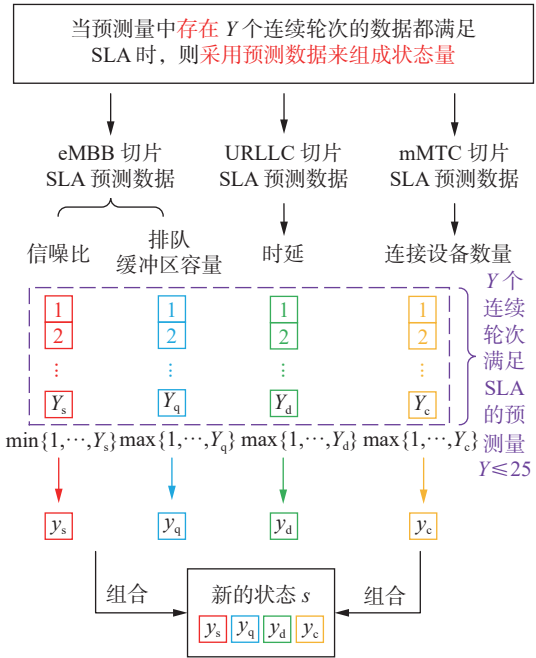


图 5 预测数据处理流程

Fig. 5 Predictive data processing procedure

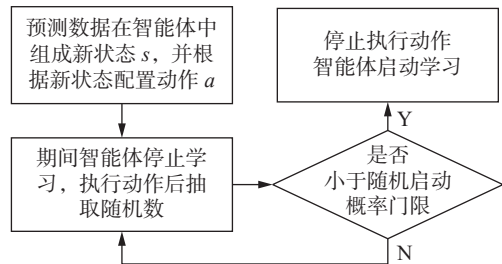


图 6 预测数据处理流程

Fig. 6 The predictive data processing procedure

当随机数小于随机启动概率门限, 或者针对没有  $Y$  个连续轮次的数据都满足 SLA 的情况, 智能体将不采用预测数据组成新状态量。而是在每个轮次开始前, 先对系统环境进行观测获取观测状态, 再基于观测状态进行强化学习配置动作, 随后进入下一个轮次, 依次循环。

### 2.4 问题定义

将 5G 切片时频资源分配问题采用马尔可夫决策过程 (Markov decision process, MDP) 建模, 包含以下数学定义:  $(S, A, P_r, R, \gamma)$ , 目标是寻找一个合适的策略, 使得分配资源的数量最小化。其中  $S$  表示所有可能的状态集合, 即状态空间;  $A$  表示智能体可以选择执行的所有可能行为或动作的集合;  $P_r(s, s') = P(s_{n+1} = s' | s_n = s, a_n = a)$  表示从当前状态  $s$  执行动作  $a$  后转移到下一个状态  $s'$  的概率分

布;  $R(s, s')$  表示在执行动作  $a$  后, 智能体从状态  $s$  转移到  $s'$  时获得的奖励;  $\gamma$  是一个介于 0~1 的数值, 表示折扣因子, 决定了对未来奖励的重视程度。在 5G 切片时频资源分配问题中, 将重点关注状态  $S$ 、动作  $A$  以及奖励  $R$ 。

#### 2.4.1 状态空间

由于 Stacked-LSTM-GKBRL 算法是根据当前时刻的观测值来计算并做出最佳动作, 从而在尽可能有效地分配 RBs 前提下, 同时保障实现终端设备的 SLA。因此状态  $s_n$  需要反映出当前网络中各切片指标状态。对于每一个切片  $i \in B$ , 在每个轮次  $n = 1, 2, \dots$ , 将  $s_{n(i)}$  定义为智能体在轮次  $n$  接收切片  $i$  的观测值。  $s_n = (s_{n(1)}, s_{n(2)}, \dots, s_{n(B)})$  表示为系统观测值或者状态空间值, 即第  $n$  轮次每个切片观测值的集合。  $s_{n(i)}$  主要包含信噪比、排队缓冲区容量、时延和连接设备数量等观测变量。  $s_{n(i)}$  在轮次  $n$  经过预处理后, 反映当前网络中各切片的流量请求。

#### 2.4.2 动作空间

本文对于切片资源分配的问题, 将动作定义为给每个业务类型切片分配对应的时频资源。因此, 将  $a_{n(i)}$  表示在轮次  $n$  智能体分配给切片  $i$  的 RBs 数量。  $a_n = (a_{n(1)}, a_{n(2)}, \dots, a_{n(B)})$  表示在轮次  $n$  分配给每个切片的时频资源集合, 即动作空间值。

#### 2.4.3 奖励函数

为衡量切片的 SLA 违反情况, 首先设定指示函数  $I_{(i)}$  包含 SLA 中定义的关键指标, 每个轮次向智能体反馈各切片的 SLA 实现情况。  $I_{n(i)}(s_{n(i)}, a_{n(i)}) = 1$  表示在轮次  $n$  违反了第  $i$  个切片的 SLA; 若为 0, 则表示没有违反 SLA。当出现 2.3 节中讨论到的 SLA 长时间未被违反的情况时, 智能体可能陷入局部最优的情况。因此仍需引入另一个指示函数  $I_Y$ ,  $I_Y = 1$  表示  $Y$  个连续轮次的数据都满足 SLA; 反之  $I_Y = 0$ 。对奖励函数定义为

$$R = \gamma I_{n(i)}(s_{n(i)}, a_{n(i)}) + a_{n(i)} + \varepsilon I_Y \quad (3)$$

$R$  表示在轮次  $n$  观测动作对的奖励, 其中  $\gamma$  和  $\varepsilon$  分别为折扣因子。如果违反了切片的 SLA, 说明当前轮次下配置的动作无法满足业务需求, 需要分配更多的时频资源。而如果满足了切片的 SLA, 说明当前轮次下配置的动作已满足业务需求, 进而尝试减少分配的资源数量, 探索准最优的时频资源分配策略, 从而达到节约资源的目的。本算法通过适当减少奖励, 促使系统分配更少的时频资源, 以此激动智能体向外探索, 逐步接近准最优的时频资源分配策略, 最终达到节约资源的目的。

#### 2.4.4 优化目标

由于指示函数会直接影响奖励函数, 为保证指示函数的准确性, 将假设函数  $j_{n(i)}$  定义为: 在给出上一轮次结束时收到的观测值  $s_{n-1(i)}$  的基础上, 来假设判断下一轮次的  $a_{n(i)}$  是否满足 SLA。但由于不够准确的假设可能会导致下一轮次的  $a_{n(i)}$  会过多违反 SLA, 所以还需要设置误差函数。将误差函数  $e_{n(i)}$  定义为: 由  $j_{n(i)}$  在观测-动作对  $(s_{n(i)}, a_{n(i)})$  上给出的假设是没有违反 SLA 而实际却违反了 SLA 的概率, 即  $j_{n(i)}$  误判的概率。因此误差函数以差错概率  $\delta$  为界, 为寻找最优的资源分配策略, 将式 (1) 中的目标函数优化为

$$\begin{aligned} \min \sum_{i=1}^B a_{n(i)}, \text{ for } i \in B, \\ \text{s.t. } e_{n(i)}(j_{n(i)}, s_{n-1(i)}, a_{n(i)}) \leq \delta, \sum_{i \in B} a_{n(i)} \leq C \end{aligned} \quad (4)$$

#### 2.5 GKBRL 模型

关于 GKBRL 算法主要由假设模块、误差模块和决策模块组成。

##### 2.5.1 假设模块

假设模块的作用是假设每个网络切片在给定时频资源分配下是否能够满足其 SLA。在实际的检测流水线中, 面对复杂多样的环境, 简单的线性在线学习无法准确模拟真实场景中众多的非线性检测情况。为解决这一缺点, 其中一种较为有效的策略是使用核函数<sup>[18]</sup>。将假设函数  $j_n$  写成核展开式:

$$j_n(x) = \sum_{x_n \in \mathcal{X}_n} I_n k(x_n, x) \quad (5)$$

式中: 核  $k(x_n, x)$  表示新样本  $x$  和样本集中已有样本  $x_n$  之间的相似性;  $I_n$  是样本集中已有样本  $x_n$  对应各自的指示函数;  $\mathcal{X}_n$  为假设结果错误的样本集。

关于核函数  $k(x_n, x)$ , 由于高斯核适用于高度非线性的数据, 故可将  $k(x_n, x)$  写为

$$k(x_n, x) = \exp\left(-\frac{\|x_n - x\|^2}{2\sigma^2}\right) \quad (6)$$

式中:  $\sigma$  为控制核函数的宽度,  $\|x_n - x\|$  为两个向量的欧几里得距离:

$$\|x_n - x\|^2 = \sum_{k=1}^d (x_{nk} - x_k)^2 \quad (7)$$

式中:  $d$  表示向量的维度。当  $x_n$  和  $x$  非常接近时,  $k(x_n, x)$  接近 1; 当  $x_n$  和  $x$  相距较远时,  $k(x_n, x)$  接近 0。

如果假设结果发生错误则将新样本  $x_{n+1}$  添加到样本集  $\mathcal{X}_n$  中, 成为新样本集  $\mathcal{X}_{n+1}$ , 并同时更新假设函数:

$$j_{n+1}(x) = \sum_{x_{n+1} \in \mathcal{X}_{n+1}} I_{n+1} k(x_{n+1}, x) \quad (8)$$

将更新后的假设函数  $j_{n+1}$  作为智能体的下一轮次输入。

### 2.5.2 误差模块

为了评估假设模块的准确性, 给定切片  $i$  的观测值  $s_{n-1(i)}$ 。在假设能够满足切片  $i$  的 SLA 的动作中, 如果  $a$  是最小的动作, 则将动作  $a_{n(i)}$  的安全边界  $m_{n(i)}$  定义为

$$m_{n(i)} = a_{n(i)} - a \quad (9)$$

在假设满足 SLA 的动作中, 安全边界越大, 误差函数  $e_{n(i)}$  越小, 意味着出现假设错误的概率越低。若安全边界越接近 0, 其假设错误的概率越大, 越容易违反 SLA。由于误差函数  $e_{n(i)}$  定义为  $j_{n(i)}$  的误判概率, 所以只有  $j_{n(i)}$  出现假设错误时才会更新  $e_{n(i)}$ , 再根据 SLA 实际满足情况  $I_{n(i)}$  进而精确误差函数。因此, 误差函数  $e_{n+1(i)}(m)$  为

$$e_{n+1(i)}(m) = (1 - \beta)e_{n(i)}(m) + \beta I_{n(i)} \quad (10)$$

其中  $\beta$  是学习率。在每个轮次的最后输出各切片对应的误差函数  $e_{n+1}$ , 并将其作为智能体的下一轮次输入。

### 2.5.3 决策模块

决策模块接收前一轮次结束时的观测值  $s_{n-1}$ 、轮次  $n$  的假设函数集  $j_n$  和误差函数集  $e_n$ 。在预测能够满足切片  $i$  的 SLA 的动作中, 最小的动作  $a$  对应的误差函数不一定能满足小于式 (4) 中以  $\delta$  为界的差错概率, 故还需进一步寻找最优动作。即寻找能够满足以下两个条件的最小动作:

- 1) 该动作能够满足切片  $i$  的 SLA;
- 2) 该动作对应的误差函数小于差错概率  $\delta$ 。

对于每个切片, 决策模块根据各切片对应的误差函数以  $e_{n(i)} \leq \delta$  为条件, 选出各切片的最小安全边界  $m_{n(i)}$ 。同时选出最小的能够满足 SLA 的动作  $a_{n(i)}$ , 再将  $m_{n(i)}$  和  $a_{n(i)}$  相加即可得到轮次  $n$  各切片对应的最优动作  $a$ 。具体过程如算法 1 所示。

#### 算法 1 Stacked-LSTM-GKBRL

**输入** 资源块总数  $C$ , 差错概率  $\delta$ , 各切片得到的最大资源块数量  $a_{\max}$

1) 初始化假设函数集  $j_0$ , 误差函数集  $e_0$ , 起始轮次  $n_{\text{start}}$ , 结束轮次  $n_{\text{end}}$

2) 从环境中获取初始状态观测值  $s_0$

3) for 轮次  $n = 1, 2, \dots, N$  do

4) if  $n|200 = 0$

5) Stacked-LSTM 根据历史 200 个轮次观测值对未来 200 个轮次进行预测, 并将预测结果输入至智能体中

6) 智能体根据预测结果组合新状态  $s$ , 并更新起始轮次  $n_{\text{start}}$ 、结束轮次  $n_{\text{end}}$

7) if  $n = n_{\text{start}}$

8) 决策模块根据新状态  $s$ 、前一轮次的假设函数  $j_n$ 、误差函数  $e_n$  生成并执行动作  $a_n$ , 记录观测值  $s_n$ , 获得奖励函数  $R_n$

9) continue

10) if  $n_{\text{start}} < n \leq n_{\text{end}}$

11) 智能体继续执行动作  $a_n$ , 记录观测值  $s_n$ , 获得奖励函数  $R_n$

12) if 随机数小于随机启动概率门限

13)  $n_{\text{end}} = 0$

14) continue

15) 更新假设函数  $j_n$ 、误差函数  $e_n$

16) 决策模块执行动作  $a_n$ , 得到并记录观测值  $s_n$ , 获得奖励函数  $R_n$

17) end for

## 3 仿真分析

为了验证 Stacked-LSTM-GKBRL 模型在预测准确性方面的优势, 本实验在 PyCharm 仿真平台中搭建相关仿真实验, 用于模拟在 5G 混合场景下进行 RAN 切片之间分配时频资源。该场景中存在一个基站, 服务于 3 种不同业务类型的切片: eMBB、URLLC 和 mMTC。网络切片环境的物理参数相关设置具体如表 1 所示。

表 1 无线网络环境仿真参数设置

Table 1 Simulation parameter configuration for wireless network environments

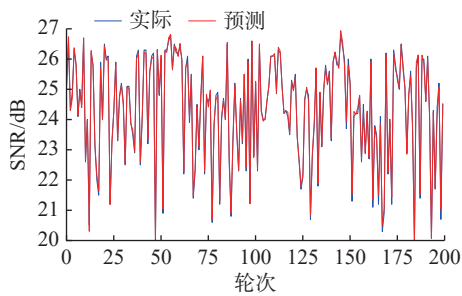
参数名称	参数值
基站发射功率/dBm	30
基站天线增益/dBi	15
载波频率/GHz	2
噪声系数/dB	9
小区半径/km	2
信道	频率选择性衰落

为了验证 Stacked-LSTM-GKBRL 模型在预测准确性与资源分配方面的优势, 设计了一个由 2 个 eMBB 切片、2 个 URLLC 切片和 1 个 mMTC 切片组成, 每个子帧最多可使用 150 个 RBs 的仿真场景。各切片业务模型具体设置如表 2 所示, 其中 mMTC 业务模型的传输周期和数据包重传次数在所示的集合中随机选择<sup>[18,29]</sup>。

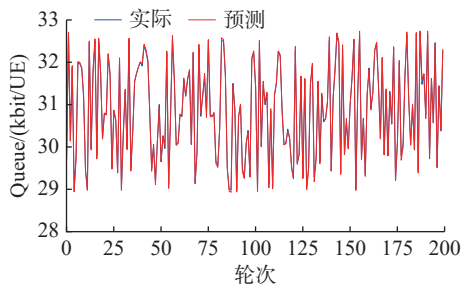
表 2 RAN 切片流量关键参数设置  
Table 2 Configuration of key parameters for RAN slicing traffic

业务模型	关键参数	数值
eMBB业务模型	用户到达过程(泊松过程平均到达率)/(用户/min)	2
	用户连接时间/s	服从指数分布, 平均值为30
	速率/(Mb/s)	0.5
	SLA需求: 信噪比/dB	不低于20
	SLA需求: 排队缓冲容量/(kbit/用户)	不超过100
URLLC业务模型	数据包大小/bits	100
	信噪比/dB	10
	可靠性/%	99.999
	SLA需求: 时延/ms	不高于5
mMTC业务模型	数据包大小/bits	100
	传输周期/s	{10,50,100,150,200,250,500,1000}
	数据包重传次数	{2,4,8,16,32,64,128}
	SLA需求: 同时连接设备数量/个	不超过1000
	随机启动概率门限 $p_{start}/\%$	4

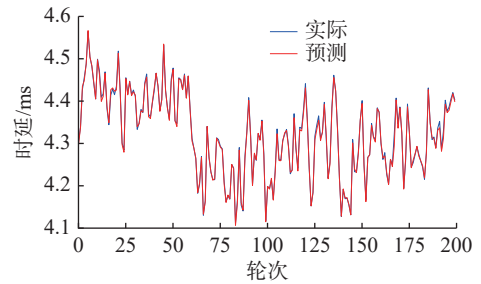
如图 7 所示, 其较为直观地显示了 Stacked-LSTM 预测模型预测所得结果与实际值的对比关系。可以明显地看到, 预测后的数据曲线与真实曲线趋势基本保持一致, 大部分情况下较为接近, 即使真实数据产生波动, Stacked-LSTM 预测模型也能大致拟合出真实数据的分布情况。



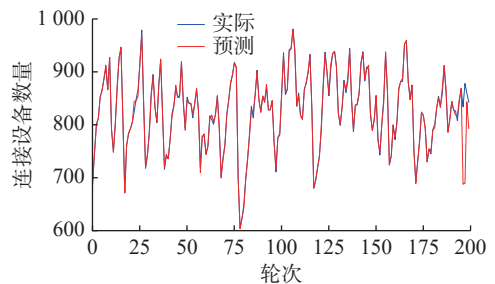
(a) 信噪比预测分布



(b) 排队缓冲容量预测分布



(c) 时延预测分布



(d) 连接设备数量预测分布

图 7 预测数据与真实数据对比  
Fig. 7 Comparison of predicted and actual data

如图 8 所示, 训练损失曲线在最初的几个 epoch 内迅速下降, 然后逐渐趋于平稳, 表明 Stacked-LSTM 预测模型正在逐步学习到数据的特征, 并在减少误差。测试损失曲线没有出现明显的上升趋势, 而是随着训练进行逐渐下降并趋于平稳,

无明显波动情况, 这说明 Stacked-LSTM 预测模型在测试集上的误差稳定, 泛化能力有所提升, 没有明显的过拟合或欠拟合情况发生。

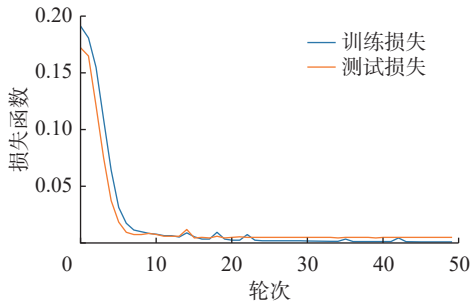


图 8 Stacked-LSTM 算法损失函数曲线  
Fig. 8 Loss curves of the Stacked-LSTM algorithm

为了对比 Stacked-LSTM 模型的预测精确性能, 设置了另外 4 种预测模型: 循环神经网络 (recurrent neural network, RNN)、门控循环单元 (gated recurrent unit, GRU)、极端梯度提升 (extreme gradient boosting, XGBoost) 和长短期记忆网络 LSTM 的对照实验。图 9 为 5 种算法对信噪比预测的残差 (预测值与真实值的差异) 的分布图所示。总体来看, Stacked-LSTM 预测模型的残差分布接近正态分布, 说明残差是随机分布的, 没有明显的系统性偏差。可以看到 Stacked-LSTM 预测模型在预测时的大部分残差主要集中在  $[-0.2, 0.2]$  dB 之间, 极端残差较少, 残差大致符合单峰的钟形曲线, 这表明 Stacked-LSTM 预测模型的预测误差是对称的、偏差不大。与 RNN、GRU、XGBoost 以及 LSTM 模型的残差分布相比, Stacked-LSTM 模型的残差分布较为集中, 而其余 4 种模型的分布相对分散, 表明 Stacked-LSTM 预测模型的偏差较小, 预测性能表现较为良好。

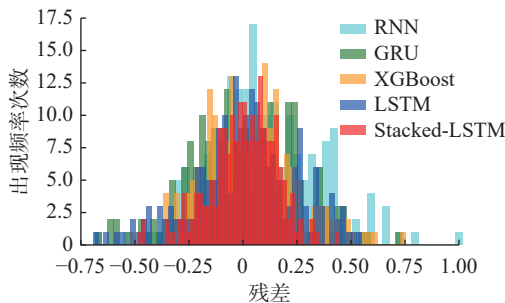


图 9 预测残差分布  
Fig. 9 Residual distribution of the prediction results

图 10 为 Stacked-LSTM、RNN、XGBoost、GRU 和 LSTM 5 种预测算法关于信噪比预测结果的均方误差图。可以看到 Stacked-LSTM 预测模型的 MSE 值低于 RNN、XGBoost、GRU 和 LSTM 这 5 种算法。并且随着轮次增加, Stacked-LSTM 预

测模型的 MSE 值起伏较为稳定, 并没有明显波动。说明训练后的 Stacked-LSTM 预测模型在实际预测过程中表现较为良好, 没有出现明显的预测偏差。在 150~200 轮次中 Stacked-LSTM 预测模型的 MSE 值略微增加, 这是由于 Stacked-LSTM 预测模型是在前一个预测结果的基础上继续预测的, 说明预测的误差不可避免地随着轮次的增加而在逐步积累。

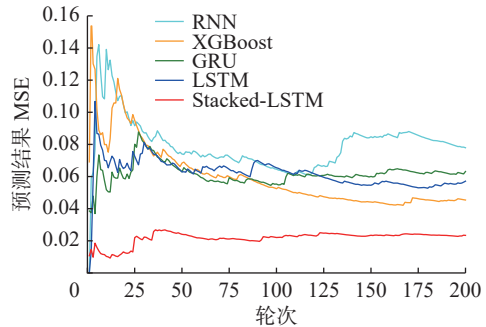


图 10 预测结果 MSE  
Fig. 10 MSE curves of prediction results

为了充分评估 Stacked-LSTM-GKBRL 模型各方面的性能, 本文将其与 4 种经典强化学习算法, 即 TD3 (twin delayed deep deterministic policy gradient)、NAF (normalized advantage function)、SAC (soft actor critic) 以及 A2C (synchronous advantage actor critic) 算法, 还有未采取预测的 GKBRL 进行比较。同时还设置了  $Y \leq 25$ 、 $Y \leq 50$  和  $Y \leq 80$  共 3 种对照情况。

在图 11 的资源分配图中, 对于每个算法绘制了 90% 置信水平的平均值曲线以及置信区间。可以看出在另外的 4 种算法中, 关于算法收敛速度性能上, A2C 和 NAF 算法收敛速度较快; 而 NAF、GKBRL 和 Stacked-LSTM-GKBRL 系列曲线在前 3000 步初始轮次中呈现上升趋势, 随后在后续轮次呈现逐步收敛; TD3 曲线更是到 6000 步轮次左右才逐渐收敛。在资源分配上, NAF 和 A2C 曲线基本稳定在 120 个 RBs 左右, 体现为在早期轮次过度配置资源; 而 TD3 和 SAC 曲线则是在 120~145 个 RBs 区间内逐渐上升, 在尽量满足 SLA 的条件下分配的资源趋近于上限 150 个 RBs, 体现为过度地配置了资源。相比之下, 在前 3000 步的初始轮次中, GKBRL 和 Stacked-LSTM-GKBRL 系列曲线快速上升, 随后算法呈现逐步收敛趋势, 资源分配趋于稳定。Stacked-LSTM-GKBRL 能够防止在早期轮次过度配置资源, 从而优化了资源利用。相较之下, 作为对照的算法收敛速度虽然相差不大, 但其在资源分配上浪费了较

多资源,资源分配精度性能表现远远不如本文所提算法。

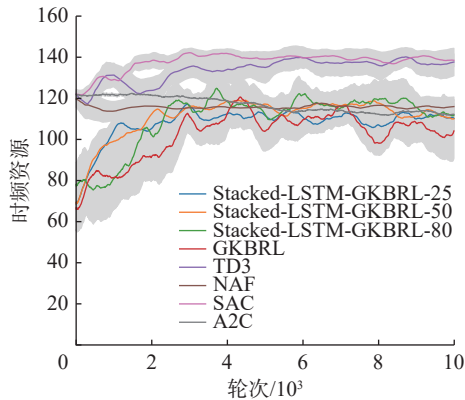


图 11 4种算法资源分配

Fig. 11 Resource allocation results of four algorithms

从 GKBRL 和 Stacked-LSTM-GKBRL 系列曲线中可以看出, GKBRL 算法总体分配的时频资源相较于 Stacked-LSTM-GKBRL 系列算法要少,而且 Stacked-LSTM-GKBRL 总体分配的资源随着  $Y$  的数量增加而增加,在资源分配精度性能上也随之下降。这是由于结合了 Stacked-LSTM 预测模型后,为了保证满足 SLA 而需要预留额外的资源,在资源分配精度性能和满足 SLA 性能中进行了折中。同时由于选取轮次的区间增加,需要考虑更多相对性能较差的样本,从而导致总体需要分配的资源增加。但图 11 中 Stacked-LSTM-GKBRL 系列曲线距离较近,说明所需要预留的额外资源并不是特别多,并没有过多占用资源。此外, Stacked-LSTM-GKBRL-50 和 Stacked-LSTM-GKBRL-80 的曲线比较接近,说明当  $Y$  大于 50 之后,即使  $Y$  区间数量再多,需要分配的资源也没有明显的增加。

图 12 为累计的 SLA 违反情况。从图中可以明显地看出 NAF、TD3、SAC 和 A2C 曲线显著高于 Stacked-LSTM-GKBRL 系列曲线,说明 Stacked-LSTM-GKBRL 在避免违反 SLA 上表现优异。同时, Stacked-LSTM-GKBRL 系列曲线总体上低于 GKBRL 曲线,说明使用 Stacked-LSTM 预测模型后能够有效地降低违反 SLA 的情况,性能得到进一步改进。

此外,由于使用 Stacked-LSTM 预测模型后,智能体不需要频繁的学习,不仅节约了算力资源、降低了算法的复杂度,在 SLA 满足率上还依然能够接近 GKBRL 算法,甚至在某些轮次上比 GKBRL 算法还要表现优异。结合资源分配与 SLA 违规数量两个指标,  $Y$  取值不大于 50 为相对最优解。

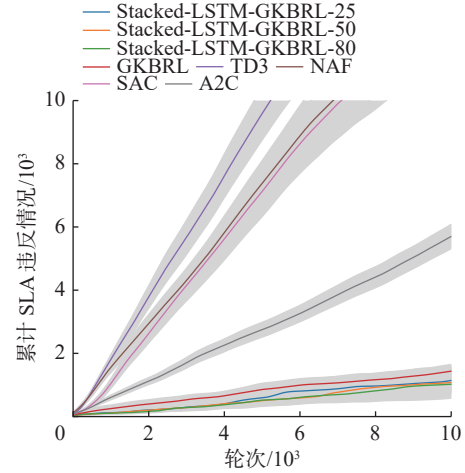


图 12 累计 SLA 违反情况

Fig. 12 Cumulative SLA violations

## 4 结束语

本文所提出的基于 Stacked-LSTM 与 GKBRL 相结合的智能切片时频资源优化分配策略,能够有效应对载波通信模块检测流水线中网络状态快速变化带来的资源调度挑战。通过引入时间序列预测提升状态感知能力,并结合高斯核强化学习优化资源分配决策,实现了在保障 SLA 指标的前提下对时频资源的高效利用。实验结果验证了该算法在预测精度和分配性能方面的优势,为 5G 资源管理提供了一种具有实用价值的解决方案。

## 参考文献:

- [1] 殷珉, 沈航, 王天荆, 等. 基于分层联邦学习的无人机小基站 RAN 切片方法[J]. 电子学报, 2023, 51(7): 1774–1780.  
YIN Min, SHEN Hang, WANG Tianjing, et al. Hierarchical federated learning-based RAN slicing for drone-small-cells[J]. Acta electronica sinica, 2023, 51(7): 1774–1780.
- [2] GUNDALL M, STRUFE M, SCHOTTEN H D, et al. Introduction of a 5G-enabled architecture for the realization of industry 4.0 use cases[J]. IEEE access, 2021, 9: 25508–25521.
- [3] WANG Haozhe, WU Yulei, MIN Geyong, et al. A graph neural network-based digital twin for network slicing management[J]. IEEE transactions on industrial informatics, 2022, 18(2): 1367–1376.
- [4] MAHMOOD A, BELTRAMELLI L, FAKHRUL ABEDIN S, et al. Industrial IoT in 5G-and-beyond networks: vision, architecture, and design trends[J]. IEEE transactions on industrial informatics, 2022, 18(6):

- 4122–4137.
- [5] 王再见, 谷慧敏. 基于联合优化的网络切片资源分配策略[J]. *通信学报*, 2023, 44(5): 234–245.  
WANG Zaijian, GU Huimin. Network slicing resource allocation strategy based on joint optimization[J]. *Journal on communications*, 2023, 44(5): 234–245.
- [6] THIRUVASAGAM P K, CHAKRABORTY A, MURTHY C S R. Resilient and latency-aware orchestration of network slices using multi-connectivity in MEC-enabled 5G networks[J]. *IEEE transactions on network and service management*, 2021, 18(3): 2502–2514.
- [7] DATAR M, ALTMAN E, LE CADRE H. Strategic resource pricing and allocation in a 5G network slicing Stackelberg game[J]. *IEEE transactions on network and service management*, 2023, 20(1): 502–520.
- [8] AYEPAH-MENSAH D, SUN Guolin, BOATENG G O, et al. Blockchain-enabled federated learning-based resource allocation and trading for network slicing in 5G[J]. *IEEE/ACM transactions on networking*, 2024, 32(1): 654–669.
- [9] AZIMI Y, YOUSEFI S, KALBKHANI H, et al. Applications of machine learning in resource management for RAN-slicing in 5G and beyond networks: a survey[J]. *IEEE access*, 2022, 10: 106581–106612.
- [10] 陈智雄, 詹学滋, 左嘉烁. 基于深度强化学习的电力线与无线双模通信 MAC 层接入算法[J]. *智能系统学报*, 2025, 20(2): 344–354.  
CHEN Zhixiong, ZHAN Xuezi, ZUO Jiashuo. Adaptive MAC layer access algorithm for power line and wireless dualmode communication based on deep reinforcement learning[J]. *CAAI transactions on intelligent systems*, 2025, 20(2): 344–354.
- [11] YARKINA N, GAYDAMAKA A, MOLTCHANOV D, et al. Performance assessment of an ITU-T compliant machine learning enhancements for 5G RAN network slicing[J]. *IEEE transactions on mobile computing*, 2024, 23(1): 719–736.
- [12] DONG Tianjian, QI Qi, WANG Jingyu, et al. Standing on the shoulders of giants: cross-slice federated meta learning for resource orchestration to cold-start slice[J]. *ACM transactions on networking*, 2023, 31(2): 828–845.
- [13] HU Tianlun, LIAO Qi, LIU Qiang, et al. Information bottleneck-based domain adaptation for hybrid deep learning in scalable network slicing[J]. *IEEE transactions on machine learning in communications and networking*, 2024, 2: 1642–1660.
- [14] KRISHNA M B, LORENZ P. Deterministic network slice instance policy for intra and inter slice resource management in 5G[J]. *IEEE transactions on vehicular technology*, 2025, 74(3): 4904–4916.
- [15] ZHOU Hao, EROL-KANTARCI M, VINCENT POOR H. Learning from peers: deep transfer reinforcement learning for joint radio and cache resource allocation in 5G RAN slicing[J]. *IEEE transactions on cognitive communications and networking*, 2022, 8(4): 1925–1941.
- [16] CHANG Xiaolei, JI Tian, ZHU Runsu, et al. Toward an efficient and dynamic allocation of radio access network slicing resources for 5G era[J]. *IEEE access*, 2023, 11: 95037–95050.
- [17] 李冲冲. 基于模型的强化学习中随机算法研究[D]. 北京: 北京交通大学, 2024.  
LI Chongchong. Research on random algorithm in model-based reinforcement learning[D]. Beijing: Beijing Jiaotong University, 2024.
- [18] ALCARAZ J J, LOSILLA F, ZANELLA A, et al. Model-based reinforcement learning with kernels for resource allocation in RAN slices[J]. *IEEE transactions on wireless communications*, 2023, 22(1): 486–501.
- [19] KATTEPUR A, DAVID S, MOHALIK S K. Model-based reinforcement learning for router port queue configurations[J]. *Intelligent and converged networks*, 2021, 2(3): 177–197.
- [20] CHEN Yuxuan, LI Rongpeng, YU Xiaoxue, et al. Adaptive layer splitting for wireless large language model inference in edge computing: a model-based reinforcement learning approach[J]. *Frontiers of information technology & electronic engineering*, 2025, 26(2): 278–292.
- [21] 王迪, 张震, 王洋, 等. 一种基于模型可信度的强化学习方法[J/OL]. *控制工程*, 1–9[2025–06–16].  
WANG Di, ZHANG Zhen, WANG Yang, et al. A reinforcement learning method based on model credibility [J/OL]. *Control engineering of China*, 1–9[2025–06–16].
- [22] YU Peng, YANG Mo, XIONG Ao, et al. Intelligent-driven green resource allocation for industrial Internet of Things in 5G heterogeneous networks[J]. *IEEE transactions on industrial informatics*, 2022, 18(1): 520–530.
- [23] FERREIRA D, BRAGA REIS A, SENNA C, et al. A forecasting approach to improve control and management for 5G networks[J]. *IEEE transactions on network and service management*, 2021, 18(2): 1817–1831.
- [24] 姜华, 杨家伟, 黄巍, 等. 基于深度强化学习的 D2D 辅助 MEC 网络资源分配算法[J]. *电力信息与通信技术*, 2023, 21(7): 51–58.  
JIANG Hua, YANG Jiawei, HUANG Wei, et al. A D2D-

- assisted MEC network resource allocation algorithm based on deep reinforcement learning[J]. *Electric power information and communication technology*, 2023, 21(7): 51–58.
- [25] MARTIN-PEREZ J, KONDEPU K, DE VLEESCHAUWER D, et al. Dimensioning V2N services in 5G networks through forecast-based scaling[J]. *IEEE access*, 2022, 10: 9587–9602.
- [26] 陈志鹏, 张勇, 高海荣, 等. 隐私保护下融合联邦学习和 LSTM 的少数数据综合能源多元负荷预测[J]. *智能系统学报*, 2024, 19(3): 565–574.
- CHEN Zhipeng, ZHANG Yong, GAO Hairong, et al. Integrated energy multivariate load forecasting combining federated learning with LSTM in privacy-protected and low-data environments[J]. *CAAI transactions on intelligent systems*, 2024, 19(3): 565–574.
- [27] WU Qiong, CHEN Xu, ZHOU Zhi, et al. Deep reinforcement learning with spatio-temporal traffic forecasting for data-driven base station sleep control[J]. *IEEE/ACM transactions on networking*, 2021, 29(2): 935–948.
- [28] MENDOZA J, DE-LA-BANDERA I, PALACIOS D, et al. Forecasting framework for mobile networks based on automatic feature selection[J]. *IEEE transactions on network and service management*, 2023, 20(2): 974–984.
- [29] 赵晨, 张铖, 黄永明. 基于流量感知的无线接入网智能切片资源分配方法研究[J]. *信号处理*, 2024(4): 719–732.
- ZHAO Chen, ZHANG Cheng, HUANG Yongming. Research on traffic-aware intelligent slicing resource allocation for radio access network[J]. *Journal of signal processing*, 2024(4): 719–732.
- [30] 邓翠艳, 齐小刚. 一种注意力机制 LSTM 的 5G 网络地铁节电方法[J]. *智能系统学报*, 2024, 19(5): 1309–1318.
- DENG Cuiyan, QI Xiaogang. 5G network subway power-saving method based on attention mechanism LSTM[J]. *CAAI transactions on intelligent systems*, 2024, 19(5): 1309–1318.

#### 作者简介:



陈智雄, 副教授, 主要研究方向为电力物联网、电力线通信。主持国家自然科学基金项目、河北省自然科学基金等科研项目 10 余项。发表学术论文 30 余篇, 获得国家发明专利授权 6 项。E-mail: [zxchen@ncepu.edu.cn](mailto:zxchen@ncepu.edu.cn)。



谢宇鹏, 硕士研究生, 主要研究方向为 5G 切片。E-mail: [2751263429@qq.com](mailto:2751263429@qq.com)。



郭以贺, 讲师, 博士研究生, 主要研究方向为智能配电网、中压电力线通信。E-mail: [yihe\\_guo@163.com](mailto:yihe_guo@163.com)。