

基于位置预测模型的空天地一体化网络切换算法

谢健骊, 陈龙, 张泽鹏, 李翠然

(兰州交通大学电子与信息工程学院, 甘肃 兰州 730070)

摘要: 针对 6G 空天地一体化网络 (SAGIN) 中网络环境动态变化和用户终端移动性增强导致的终端切换频繁、网络负载不均衡问题, 提出了一种基于终端位置预测模型的 SAGIN 切换算法。该算法构建了基于麻雀搜索策略优化的长短期记忆 (LSTM) 网络终端位置预测模型, 提升了终端位置预测精度, 解决了网络切换时机不合理问题。基于此模型, 将 SAGIN 选择问题建模为马尔可夫决策过程, 设计以服务质量 (QoS) 需求、切换代价和网络负载均衡表征的网络切换算法效用函数, 采用分布式深度 Q 网络 (D-DQN) 选择能够实现长期目标最大化的网络节点执行切换。与基于 Q 学习 (Q-Learning)、双深度 Q 网络 (DDQN) 和竞争双深度 Q 网络 (D3QN) 的网络切换算法相比, 所提算法在降低切换时延与切换次数、提升网络吞吐量等方面性能较优, 验证了所提算法的有效性。

关键词: 空天地一体化网络; 网络切换; 效用函数; 长短期记忆网络; 分布式深度 Q 网络

中图分类号: TN929.5

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2024266

Handover algorithm for space-air-ground integrated network based on location prediction model

XIE Jianli, CHEN Long, ZHANG Zepeng, LI Cuiran

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: To address the issues of frequent handovers and network load imbalance caused by dynamic changes in the network environment and enhanced mobility of user terminals in the 6G space-air-ground integrated network (SAGIN), a handover algorithm for SAGIN based on a terminal location prediction model was proposed. The algorithm constructed a long short-term memory (LSTM) network terminal location prediction model optimized based on the sparrow search strategy, improving the accuracy of terminal location prediction and resolving the issue of unreasonable handover timing. Based on this model, the SAGIN selection problem was modeled as a Markov decision process. A network handover algorithm utility function characterized by quality of service (QoS) requirements, handover cost, and network load balancing was designed. A distributional deep Q-network (D-DQN) was employed to select the network nodes that could maximize long-term goals for execution handover. Compared with network handover algorithms based on Q-Learning, double deep Q-network (DDQN), and dueling double deep Q-network (D3QN), the proposed algorithm performs better in terms of reducing handover delay and frequency, as well as enhancing network throughput, thereby validating the effectiveness of the proposed algorithm.

Keywords: space-air-ground integrated network, network handover, utility function, LSTM, distributional DQN

收稿日期: 2024-07-31; 修回日期: 2024-12-05

通信作者: 谢健骊, xiejl@mail.lzjtu.cn

基金项目: 国家自然科学基金资助项目 (No.62161016)

Foundation Item: The National Natural Science Foundation of China (No.62161016)

0 引言

卫星通信、高空平台通信、无人机 (UAV, unmanned aerial vehicle) 通信等技术作为地面通信网络的补充, 在增强无线网络覆盖, 提供必要的应急通信服务方面表现优异^[1]。国际电信联盟 (ITU, international telecommunication union)、第三代合作伙伴计划 (3GPP, 3rd generation partnership project) 等国际组织已经开始了星地融合通信标准化的研究工作。未来的 6G 系统将对地面网络与非地面网络进行深度融合^[2], 构建一个多层次、全方位、高动态性和高时变性的立体异构网络——空天地一体化网络^[3] (SAGIN, space-air-ground integrated network)。SAGIN 的典型应用场景如图 1 所示。这是一种网络节点和移动终端 (MT, mobile terminal) 双重移动的超密集异构网络^[4-6], 尤其在高速铁路、飞机等场景中, 高移动性的 MT 将面临更加频繁和复杂的网络切换问题^[7], 这不仅会影响用户的通信服务体验, 还可能导致通信资源的浪费。为了解决这些问题, 需要设计一种合理、高效的网络切换方案。

针对异构网络中的网络切换问题, 国内外学者

已经进行了广泛的研究。现有多种网络切换算法, 如基于接收信号强度 (RSS, received signal strength)、博弈论、多属性决策、人工智能的网络切换算法等。文献[8]在考虑用户业务类型和网络状态的前提下, 提出了一种基于蜻蜓算法优化模糊神经网络的异构无线网络接入选择算法。然而, 该算法在设计过程中未充分考虑影响切换决策的多种因素, 且对业务类型的处理较为单一, 难以适应多样化的应用场景。文献[9]针对 SAGIN 中的动态切换问题, 提出了一种基于排队博弈模型的动态切换传输控制方案。该方案通过改进传统的排队博弈模型减少信息交互的次数, 优化用户体验。文献[10]针对卫星时变拓扑和 MT 移动性导致切换效率低的问题, 提出了一种基于时间演进图的实时切换预测方案, 能够有效减少不必要的切换。文献[11]针对 SAGIN 中频繁被动组切换问题, 提出了一种基于进化博弈的垂直切换方案, 将低轨道 (LEO, low earth orbit) 卫星作为中央控制器, 并根据服务质量 (QoS, quality of service) 计算平均效用, 以保证服务的连续性。文献[12]针对 SAGIN 场景中的网络选择问题, 提出了一种基于进化博弈和深度确定性策略梯度的网络选择算法, 以提高资源利用率和实现负载

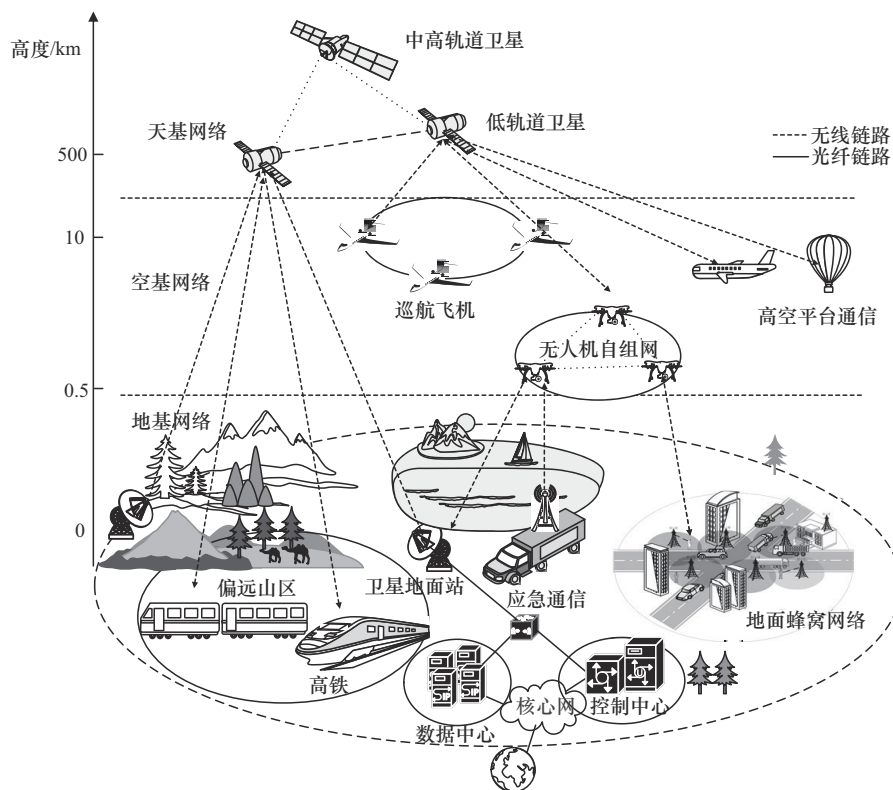


图1 SAGIN的典型应用场景

均衡。然而, 该类算法在综合考虑不同业务需求、网络状态和用户偏好等多种因素时, 面临着准确建模和解决 SAGIN 中的多方博弈问题的显著难题。文献[13]针对超密集异构网络中变速 MT 导致频繁切换的问题, 提出了一种终端移动轨迹预测的网络切换算法。该算法利用模糊核聚类和长短期记忆 (LSTM, long short-term memory) 网络进行轨迹预测, 并结合帝企鹅算法优化网络选择。文献[14]提出了一种基于区间二型模糊神经网络的垂直切换算法, 以弥补传统网络切换算法在描述网络状态随机性和模糊性方面的不足, 从而导致网络性能下降的问题。文献[15]针对卫星网络的高动态特性, 提出了一种基于用户体验的智能切换算法。该算法首先根据空间关系耦合模型预测 MT 与卫星的相对运动模式。然后利用强化学习 (RL, reinforcement learning) 算法求解多目标优化问题, 实现候选卫星的最优切换决策。然而, 如果网络节点和 MT 数量增加, RL 的动作空间和状态空间将变得更大, 从而影响算法的运行速度和效率。文献[16]提出了一种基于条件切换的切换策略, 以增强基于非地面网络中业务的连续性。然而, 该切换策略与多数现有切换策略类似, 假设用户终端的移动速度相较于 LEO 卫星可以忽略不计, 在 SAGIN 这类超密集异构网络中, 若终端移动性增强, 这一假设可能导致切换决策不准确。

上述文献主要针对地面异构网络或卫星网络场景, 基于不同的切换策略提高网络性能, 改善通信 QoS, 这为本文研究 SAGIN 融合网络场景的网络切换问题提供了很好的借鉴。然而, 这些研究存在一些局限性, 对业务类型的处理较为单一, 难以涵盖多样化的应用场景, 且未能充分体现 SAGIN 的独特性。此外, 现有算法在应对复杂动态环境时适应性不足, 难以有效应对网络环境的动态变化, 同时, 许多研究还忽视了用户在实际应用场景中的移动特性。因此, 面向具有高动态和超密集特性的 SAGIN 异构网络环境, 本文采用改进的 LSTM 网络终端位置预测模型对网络切换触发机制进行优化, 进而提出了一种综合考虑 MT 和网络节点移动性、无线链路质量、终端业务 QoS 需求以及用户偏好的 SAGIN 切换方案。本文主要贡献总结如下。

1) 针对 SAGIN 场景中网络环境动态变化和 MT

移动性增强导致 MT 频繁切换的问题, 提出了一种综合的 SAGIN 切换系统模型, MT 可以在天基、空基、地基 3 种网络中执行切换决策, 保证用户的通信 QoS。

2) 采用基于麻雀搜索算法 (SSA, sparrow search algorithm) 的智能优化策略改进 LSTM 网络终端位置预测模型, 提高了终端位置的预测精度。根据 MT 的预测位置估计未来时刻的链路质量, 预测未来的切换需求, 并设计网络切换触发机制, 高精度的终端位置预测使网络切换触发时机趋于合理化。

3) 将 SAGIN 选择问题建模为马尔可夫决策过程, 并提出了一种基于分布式深度 Q 网络 (D-DQN, distributional deep Q-network) 的网络选择算法对其进行求解。该算法综合考虑了 MT 的业务类型、不同业务的 QoS 需求、网络状态以及用户的偏好程度, 并基于效用函数理论设计了综合奖励函数, 在保证用户 QoS 需求的同时降低了切换次数和切换时延, 提升了网络性能。

1 系统模型

1.1 切换模型

SAGIN 切换模型如图 2 所示。

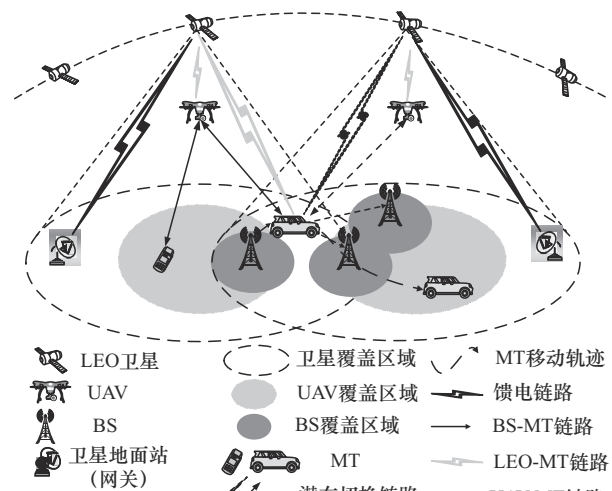


图 2 SAGIN 切换模型

图 2 中有 3 类网络节点, LEO 卫星代表天基网络节点, UAV 代表空基网络节点, 地面节点由 5G 蜂窝网络基站 (BS, base station) 和具有不同移动速度的 MT 构成。假设 SAGIN 覆盖区域内有 N 个网络节点, 记为 $D = \{D_1, D_2, \dots, D_n, \dots, D_N\}$, M 个移动终端记为 $L = \{L_1, L_2, \dots, L_m, \dots, L_M\}$ 。LEO 卫星沿固定星历轨

迹运动,当运动至MT可见范围内时,即LEO卫星与MT之间的仰角大于最小通信仰角时,可以为MT提供通信服务。每个MT可能被多个网络节点所覆盖,但MT在每个时刻只能由一个网络节点对其提供通信服务。

1.2 切换决策参数

由于SAGIN场景中网络环境的动态特性和MT的移动性,使无线链路状态呈现出多样性,每条无线链路的效用均不相同。为便于理解,在介绍具体的算法之前先给出网络切换决策参数的相关说明。

1) 载噪比 (CNR, carrier to noise ratio)

载噪比指载波信号功率与噪声信号功率之比,是衡量无线通信链路质量和卫星通信链路质量的关键参数之一。

$$\text{CNR}_{n,m,t} = \frac{P_n^{\text{tr}} G_n^{\text{tr}} G_m^{\text{rc}}}{\text{PL}_{n,m,t} k_B T_e B_n} \quad (1)$$

以分贝为单位,式(1)可由式(2)表示。

$$[\text{CNR}_{n,m,t}] = [P_n^{\text{tr}}] + [G_n^{\text{tr}}] + [G_m^{\text{rc}}] - [k_B] - [T_e] - [B_n] - [\text{PL}_{n,m,t}] \quad (2)$$

其中, $[\cdot]$ 表示变量的分贝形式, $[P_n^{\text{tr}}]$ 表示网络节点 n 的发射功率, $[G_n^{\text{tr}}]$ 表示发射机天线增益, $[G_m^{\text{rc}}]$ 表示接收机天线增益, $[k_B]$ 表示玻尔兹曼常数, $k_B = 1.38 \times 10^{-23} \text{ J/K} = -228.6 \text{ dBW}/(\text{K} \cdot \text{Hz})$, $[T_e]$ 表示接收系统的等效噪声温度, $[B_n]$ 表示网络节点 n 提供的总带宽, $[\text{PL}_{n,m,t}]$ 表示在 t 时刻网络节点 n 到终端 m (记为 MT_m)的链路传输损耗^[10]。

2) 传输速率 Rb

由香农公式可得,LEO卫星、UAV和地面网络的传输速率可表示为

$$\text{Rb}_{n,m,t} = B_{n,m} \log(1 + \text{CNR}_{n,m,t}) \quad (3)$$

其中, $B_{n,m}$ 表示 MT_m 接入网络节点 n 后分配给其的带宽, $B_{n,m} = \frac{B_n}{M_{n,t}}$,其中 $M_{n,t}$ 表示 t 时刻接入网络节点 n 的MT数量。各网络节点的带宽资源由该网络节点服务的所有MT共享,如果在同一时间内接入的MT数量增多,可能会引起网络阻塞。

3) 时延 D

对于LEO卫星和UAV网络节点,时延 D 主要考虑数据的发送时延和传播时延,可表示为

$$D_{n,m,t} = \frac{P_a}{\text{Rb}_{n,m,t}} + \frac{d_{n,m,t}}{c} \quad (4)$$

其中, P_a 表示 MT_m 传输的数据包大小, $d_{n,m,t}$ 表示从LEO卫星网络节点或UAV网络节点到 MT_m 的距离,

c 表示光速。

对于地面网络节点,时延 D 仅考虑数据的发送时延,可表示为

$$D_{n,m,t} = \frac{P_a}{\text{Rb}_{n,m,t}} \quad (5)$$

4) 误码率 (BER, bit error rate)

通信链路的误码率与信号的调制方式有关,为了简化计算,3种网络均采用二进制相移键控(BPSK, binary phase shift keying)调制方式。则BER可表示为

$$\text{BER}_{n,m,t} = \frac{1}{2} \text{erfc} \left(\sqrt{\left(\frac{E_b}{N_0} \right)_{n,m,t}} \right) \quad (6)$$

其中, $\text{erfc}(\cdot)$ 表示互补误差函数, E_b 表示每比特信号的能量, N_0 表示噪声的功率谱密度, $\left(\frac{E_b}{N_0} \right)_{n,m,t} =$

$$\text{CNR}_{n,m,t} \frac{B_{n,m}}{\text{Rb}_{n,m,t}}$$

5) 接收信号强度 RSS

对于LEO卫星、UAV和地面网络节点, RSS可表示为

$$[\text{RSS}_{n,m,t}] = [P_n^{\text{tr}}] + [G_n^{\text{tr}}] + [G_m^{\text{rc}}] - [\text{PL}_{n,m,t}] \quad (7)$$

2 基于 SSA-LSTM 和 D-DQN 的网络切换算法

2.1 问题描述与建模

SAGIN场景中的网络切换过程主要分为3个阶段:网络发现、切换决策和切换执行。其中切换决策阶段是关键,该阶段可细分为切换触发决策和切换目标网络选择决策2个过程。为提升MT的QoS和优化网络切换过程,本文基于效用函数理论,综合考虑MT业务的QoS需求、切换代价和网络负载均衡构建综合效用函数,以最大化系统总效用为优化目标,将网络选择问题转化为混合整数非线性规划问题,同时设计了相应的算法来求解优化问题。

1) 业务 QoS 属性参数和效用函数

为适应实际通信场景中MT业务的多样化特性,以及不同业务对于QoS属性参数的差异化需求,本文为各类业务设计了个性化的效用函数。参照3GPP定义的移动通信网络QoS分类,考虑了会话类、交互类、流媒体类和后台类^[17]4种业务类型。这4种业务类型对应QoS属性参数的偏好程度如表1所示。

表1 4种业务类型对应QoS属性参数的偏好程度

业务类型	敏感	不敏感	代表业务
会话类	时延,抖动,误码率,带宽		语音通话, 视频会议
交互类	误码率,时延,带宽	抖动	网页浏览, 定位
流媒体类	带宽,抖动,误码率,时延		视频播放, 音乐播放
后台类	误码率,带宽	时延,抖动	文件下载, 电子邮件

基于表1, 选取带宽 (B)、时延 (D)、抖动 (J) 和误码率 (BER) 作为评估业务 QoS 和网络状态的决策参数, 并通过线性加权法构建了 QoS 属性参数的效用函数, 以反映不同业务 QoS 属性参数在切换决策中的重要性^[18]。业务 QoS 属性参数的效用函数 $U_{n,k,t}$ 如式(8)所示。

$$U_{n,k,t} = \sum_{k=1}^K \sum_{n=1}^N w_k^C u_{n,k,t} = \sum_{n=1}^N (w_B^C u_{n,B,t} + w_D^C u_{n,D,t} + w_J^C u_{n,J,t} + w_{BER}^C u_{n,BER,t}) \quad (8)$$

其中, n 表示候选网络集中的网络节点, $n \in D$; k 表示网络 QoS 属性参数, $k \in \{B, D, J, BER\}$; K 表示 QoS 属性参数的数量, $u_{n,k,t}$ 表示 t 时刻候选网络节点 n 的 QoS 属性参数 k 的效用函数。 w_k^C 表示网络属性参数 k 的加权因子, $0 \leq w_k^C \leq 1$, 其取值既要考虑用户的个人偏好, 也要考虑网络效益。因此, 本文联合最优最劣法^[19]和熵值法得到综合权值, 以便为网络选择提供更全面的决策支持。 w_k^C 的计算式为

$$w_k^C = \eta w_k^S + (1 - \eta) w_k^O \quad (9)$$

其中, $0 \leq \eta \leq 1$ 表示权重系数, w_k^S 表示主观权重, w_k^O 表示客观权重。

在网络选择中, 效用函数被用于量化用户对网

络属性参数的偏好程度。对于带宽、信噪比等效益型网络属性参数, 效用函数采用单调递增函数进行建模。而对于时延、抖动、误码率等成本型网络属性参数则采用单调递减函数进行建模。通过设计合理的效用函数, 能够有效映射不同类型业务对于网络属性参数的要求, 效用值越大意味着该网络属性参数更加满足 MT 的业务需求^[20-21]。

各种数学函数, 如分段函数、对数函数、指数函数、Sigmoid 函数等都用来构建网络属性参数的效用模型。基于表1的分析, 给出不同网络属性参数的效用函数及其参数设置, 如表2所示。对于带宽效用模型, 会话类、交互类和后台类业务采用指数函数, 这些业务在初期对带宽较为敏感, 带宽不足时效用下降, 达到一定带宽后效用增长趋于平缓。流媒体类业务采用 Sigmoid 函数, 因其需要高带宽, 当达到阈值后效用增长趋于平缓。对于时延效用模型, 会话类业务采用 Sigmoid 函数, 该类业务对时延非常敏感, 时延的增加会导致服务质量急剧下降, 凸显了低时延对实时通信的重要性; 交互类业务采用对数函数, 该类业务时延敏感度较会话类业务低, 时延增加对用户体验影响较为缓慢; 流媒体类和后台类业务采用分段函数, 这两类业务对时延敏感度相对较低, 能适应流媒体的缓冲需求和后台任务的灵活性。对于抖动效用模型, 会话类和流媒体类业务采用 Sigmoid 函数, 这两类业务对抖动的敏感度类似, 低抖动要求严格; 交互类和后台类业务采用分段函数, 对于这两类业务, 抖动在一定范围内影响较小, 但超过限度后效用下降。对于误码率效用模型, 会话类和流媒体类业务采用分段函数, 这两类业务对于误码率的敏感度相对平滑, 但仍存在一个阈值; 交互类和后台类业务采用 Sig-

表2 不同网络属性参数效用函数及其参数设置

网络属性参数	会话类	交互类	流媒体类	后台类
带宽	指数函数 $y(x)$ $h=0.4$	指数函数 $y(x)$ $h=0.6$	Sigmoid 函数 $f(x)$ $a=0.9, b=7$	指数函数 $y(x)$ $h=0.7$
时延	Sigmoid 函数 $g(x)$ $a=0.5, b=10$	对数函数 $z(x)$ $m=-0.8, k=0.385, l=8$	分段函数 $q(x)$ $x_{\min}=15, x_{\max}=30$	分段函数 $q(x)$ $x_{\min}=30, x_{\max}=60$
抖动	Sigmoid 函数 $g(x)$ $a=1.5, b=10$	分段函数 $q(x)$ $x_{\min}=15, x_{\max}=30$	Sigmoid 函数 $g(x)$ $a=0.8, b=15$	分段函数 $q(x)$ $x_{\min}=20, x_{\max}=40$
误码率	分段函数 $q(x)$ $x_{\min}=10, x_{\max}=35$	Sigmoid 函数 $g(x)$ $a=0.5, b=12$	分段函数 $q(x)$ $x_{\min}=8, x_{\max}=25$	Sigmoid 函数 $g(x)$ $a=0.8, b=8$

moid 函数, 这两类业务对误码率非常敏感, 对数据完整性要求高。表 2 中涉及的函数表达式由式(10)和式(11)给出。

$$\begin{cases} f(x) = \frac{1}{1 + e^{-a(x-b)}} \\ g(x) = 1 - f(x) = 1 - \frac{1}{1 + e^{-a(x-b)}} \\ y(x) = 1 - e^{-hx} \\ z(x) = 1 - (m + k \ln(x+l)) \end{cases} \quad (10)$$

$$q(x) = \begin{cases} 1, & 0 \leq x \leq x_{\min} \\ \frac{x_{\max} - x}{x_{\max} - x_{\min}}, & x_{\min} < x < x_{\max} \\ 0, & x \geq x_{\max} \end{cases} \quad (11)$$

通过多样化的函数形式构建效用函数旨在准确地反映不同业务对网络属性参数的敏感度差异和偏好程度, 以满足各类业务需求的独特性。4 种业务的效用函数曲线如图 3 所示。

2) 网络切换代价效用函数

网络切换代价主要包含切换过程中产生的各种开销, 包括时间消耗和资源消耗^[12]。时间消耗定义为 MT 完成一次网络切换过程所需的时间。资源消耗主要考虑功率资源消耗和电能消耗。定义 $x_{n,m,t}$ 为决策变量, 采用二进制变量 (0 或 1) 表示在 t 时刻 MT_{*m*} 与网络节点 n 的连接关系。

$$x_{n,m,t} = \begin{cases} 1, & \text{MT}_m \text{ 与网络节点 } n \text{ 连接} \\ 0, & \text{MT}_m \text{ 与网络节点 } n \text{ 不连接} \end{cases} \quad (12)$$

此外, 令 $C_{n,t}$ 表示切换代价, 当切换的目标网络节点 n 满足业务所需 QoS 时, $C_{n,t}=0$; 否则 $C_{n,t}=\tau_{n,h}+P_{n,h}$ 。其中, $\tau_{n,h}$ 表示切换时延, $P_{n,h}$ 表示切换过程中产生的能耗, 它包括网络节点 n 的下行传输功率和网络节点 n 处的电路功率。

为了避免频繁切换, 引入切换代价效用函数 $C_{n,m,t}$, 如式(13)所示。

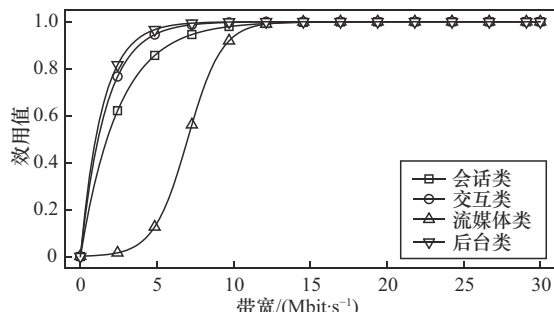
$$C_{n,m,t} = \begin{cases} 0, & x_{n,m,t} \neq x_{n,m,t-1} \\ & \text{且满足业务所需 QoS} \\ \frac{C_{n,t} - C_{\min}}{C_{\max} - C_{\min}}, & x_{n,m,t} \neq x_{n,m,t-1} \\ & \text{且不满足业务所需 QoS} \end{cases} \quad (13)$$

其中, C_{\min} 和 C_{\max} 分别表示切换代价的最小值和最大值。

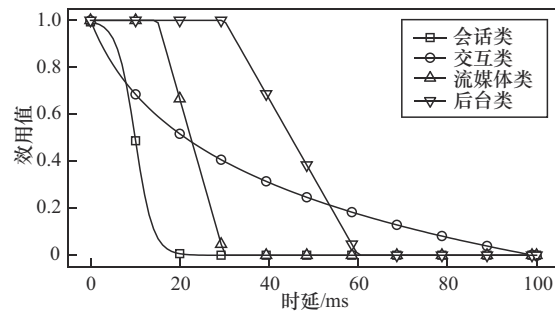
3) 负载均衡效用函数

一个网络节点可以同时为多个 MT 提供网络服务, 但每个网络节点拥有的网络资源却是有限的。当连接的 MT 数量超过网络最大承载能力时, 将导致网络阻塞, 且 MT 接入请求失败。 t 时刻网络节点 n 处的负载情况可由负载效用函数来量化^[22]。

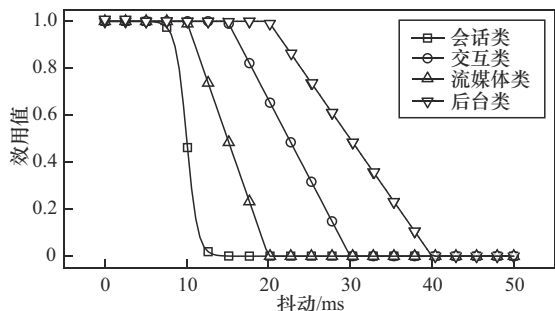
$$L_{n,t} = \frac{1}{1 + \frac{M_{n,t}}{M_n}} \quad (14)$$



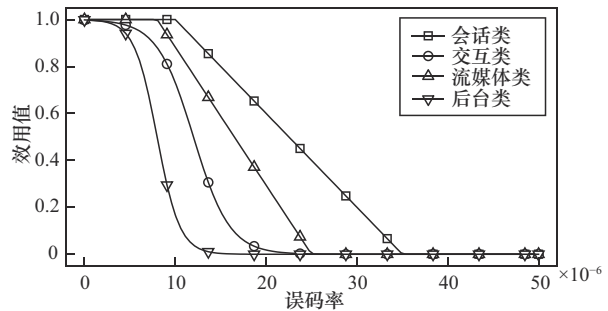
(a) 带宽效用函数曲线



(b) 时延效用函数曲线



(c) 抖动效用函数曲线



(d) 误码率效用函数曲线

图 3 4 种业务的效用函数曲线

其中, $M_{n,t}$ 表示 t 时刻接入网络节点 n 的 MT 数量, M_n 表示网络节点 n 最大可承载的 MT 数量。

本文目标是在保障用户 QoS 的同时减少不必要的切换。综合以上分析, 可将优化问题建模为综合效用函数最大化问题, 表示为

$$\begin{aligned} \text{P1: max} \quad & x_{n,m,t}(U_{n,k,t} - C_{n,m,t} + L_{n,t}) \\ \text{s.t.} \quad & C_1: x_{n,m,t} \in \{0,1\}, \forall n \in \mathcal{D}, \forall m \in \mathcal{L} \\ & C_2: 0 \leq \sum_{n=1}^N x_{n,m,t} \leq 1, \forall m \in \mathcal{L} \\ & C_3: D_{n,\text{ave}} < D_{\text{th}}, \forall n \in \mathcal{D} \\ & C_4: L_{n,\text{total}} \leq L_{\text{th}}, \forall n \in \mathcal{D} \end{aligned} \quad (15)$$

其中, C_1 表示候选网络节点 n 与 MT $_m$ 之间的连接关系, C_2 表示在 t 时刻 MT $_m$ 最多只能选择一个网络节点作为切换目标网络节点, C_3 表示平均切换时延满足最大时延阈值要求, C_4 表示候选网络节点 n 服务的 MT 数量不超过该网络负载的最大阈值。

在 t 时刻, MT $_m$ 根据式(15)选择目标函数值最大的网络节点执行切换。优化问题式(15)是一个混合整数非线性规划问题, 属于 NP-hard 问题, 难以直接获得最优解。为此, 提出了一种基于 SSA-LSTM 和 D-DQN 的网络切换算法进行求解。

2.2 基于 SSA-LSTM 的切换触发算法

切换触发决定了 MT 何时发起切换。在 SAGIN 中, 考虑 2 种可能触发切换的原因: 一是由于 MT 和网络节点的运动, MT 逐渐远离当前服务网络节点, 导致无线链路质量变差; 二是由于 MT 业务的变化, 即 MT 业务需求发生改变, 当前网络节点不能满足业务需求。本文在切换触发策略中选取 CNR 和 RSS 作为无线链路质量评估参数, 并设计了基于 SSA 优化 LSTM 网络的切换触发算法。

LSTM 网络是一种特殊的循环神经网络 (RNN, recurrent neural network)。它通过引入记忆单元、输入门、输出门和遗忘门来解决传统的 RNN 在处理长序列时遇到的“梯度消失”和“梯度爆炸”问题, 能够有效解决 RNN 无法处理长距离的依赖问题^[13]。输入门决定了当前时刻网络的输入 c_{t-1} 如何保存到记忆单元 c_t 中, 遗忘门决定了上一时刻的记忆单元 c_{t-1} 如何保存到当前时刻 c_t 中, 输出门控制当前输入数据 x_t 对当前时刻 LSTM 网络单元输出的影响, 即 LSTM 网络单元在当前时刻的最终输出 h_t 由输出门的输出 o_t 和记忆单元 c_t 共同决定。

LSTM 网络模型的性能受众多参数的影响, 包

括学习率、隐藏层层数、隐藏层的神经元数量、正则化系数、数据批处理大小等, 这些参数共同决定了模型的预测能力。然而, 在模型训练过程中主要依靠历史经验进行手动调参, 这不仅效率低下而且难以保证模型达到最优状态。因此, 采用 SSA 对 LSTM 网络模型中的学习率、LSTM 网络层神经元个数、隐藏层神经元个数和丢弃率进行优化。SSA 是一种模拟麻雀觅食行为的群体智能优化算法^[23]。利用 SSA 强大的搜索能力, 能够高效地确定神经网络模型最优超参数组合, 提高模型的预测精度。

本文以均方误差 (MSE, mean squared error) 作为 SSA 的适应度函数, 可表示为

$$\text{MSE} = \frac{1}{G} \sum_{i=1}^G (y_{i,\text{true}} - \hat{y}_{i,\text{pred}})^2 \quad (16)$$

其中, G 表示轨迹数据集中的样本个数, $y_{i,\text{true}}$ 表示真实值, $\hat{y}_{i,\text{pred}}$ 表示预测值。

利用 SSA 对 LSTM 网络模型中的 4 个超参数进行迭代寻优, 直到满足算法停止条件, 输出最优超参数组合。将得到最优超参数组合输入 LSTM 网络模型中进行 MT 位置预测, 得到 MT $_m$ 下一时刻位置。之后, 由式(17)计算 MT $_m$ 到网络节点 n 的距离, 并由式(2)和式(7)估算 MT $_m$ 下一时刻的 CNR $_{t+1}$ 和 RSS $_{t+1}$, 以此判断是否进行切换触发。

$$d_{n,m,t+1} = \sqrt{(l_{t+1,m} - l_{t+1,n})^2} \quad (17)$$

其中, $l_{t+1,m}$ 表示下一时刻 MT $_m$ 的位置坐标, $l_{t+1,n}$ 表示下一时刻网络节点 n 的位置坐标, 其中 LEO 卫星网络节点下一时刻的位置坐标可由星历数据计算得到。

当满足式(18)时, 即发现有链路质量更好的候选网络节点可用时, 将触发网络切换过程。为了提高切换触发决策的准确性和减少不必要的切换, 引入滞后余量作为额外的触发条件^[24], 切换触发策略可表示为

$$\begin{cases} \text{CNR}_{T,t+1} > \text{CNR}_{S,t+1} + \text{CNR}_{\text{Hys}} \\ \text{RSS}_{T,t+1} > \text{RSS}_{S,t+1} + \text{RSS}_{\text{Hys}} \\ \text{CNR}_{S,t+1} < \Gamma_{\text{CNR,th}} \\ \text{RSS}_{S,t+1} < \Gamma_{\text{RSS,th}} \end{cases} \quad (18)$$

其中, CNR $_{T,t+1}$ 和 RSS $_{T,t+1}$ 分别表示候选网络节点下一时刻的载噪比和接收信号强度, CNR $_{S,t+1}$ 和 RSS $_{S,t+1}$ 分别表示当前服务网络节点下一时刻的载噪比和接收信号强度, CNR $_{\text{Hys}}$ 和 RSS $_{\text{Hys}}$ 分别表示给定的载噪比和接收信号强度滞后余量, 即当前服务网

络节点和候选网络节点之间的差值, 根据 MT 位置和移动速度等设置 CNR_{Hys} 和 RSS_{Hys} 值, $\Gamma_{CNR,th}$ 和 $\Gamma_{RSS,th}$ 分别表示载噪比和接收信号强度阈值。

基于 SSA-LSTM 的切换触发算法如算法 1 所示。

算法 1 基于 SSA-LSTM 的切换触发算法

输入 MT 的轨迹数据 (如 MT_m 的位置坐标、时间戳、速度等), 网络节点 n 的位置坐标

输出 切换触发策略

- 1) 初始化参数 t 、 CNR_{Hys} 、 RSS_{Hys} 、 $\Gamma_{CNR,th}$ 、 $\Gamma_{RSS,th}$
- 2) 获取 MT 的轨迹数据, 并进行预处理 (数据清洗、归一化处理等)
- 3) 使用 LSTM 网络构建 MT 位置预测模型, 并采用 SSA 优化超参数
- 4) 将 MT 历史轨迹输入 SSA-LSTM 网络模型进行训练, 得到位置预测模型
- 5) 使用训练好的模型进行预测, 对预测结果进行反归一化处理, 输出 MT_m 下一时刻位置
- 6) 根据式 (17) 计算网络节点 n 到 MT_m 的距离 $d_{n,m,t+1}$
- 7) 根据式 (2) 和式 (7) 计算 MT_m 下一时刻的 CNR_{t+1} 和 RSS_{t+1}
- 8) if 满足式 (18) then
- 9) 激活切换触发机制, 进入网络选择阶段
- 10) else
- 11) $t=t+1$, 返回步骤 5)
- 12) end if

2.3 基于 D-DQN 的网络选择算法

根据算法 1, 当满足切换触发条件时, 会激活切换触发机制进入网络选择阶段。在 SAGIN 环境中, 采用传统的 RL 算法 (如 Q-Learning、Sarsa 等) 求解网络选择问题会面临状态空间和动作空间过大的问题, 导致算法计算量庞大, 影响算法的计算效率和收敛速度。因此, 本文采用深度强化学习 (DRL, deep reinforcement learning) 方法求解此问题。目前, 深度 Q 网络 (DQN, deep Q-network) 在基于 DRL 的网络选择领域中应用广泛。其中, 基于分布式的 DQN 性能出色, 本节将基于 D-DQN 算法对 SAGIN 选择问题进行研究。D-DQN 是一种基于价值分布的 DRL 算法。通过考虑整个奖励的分布, 而非单一的期望值, 可有效确保覆盖所有动作的价值可能性, 能够更好地适应动态、复杂的 SAGIN 选择环境, 提高网络性能。

2.3.1 D-DQN 算法原理

Q-Learning 算法是一种典型的基于值函数的 RL 算法, 它通过智能体与环境之间的不断交互学习来实现智能体累积期望奖励值达到最大化的目标, 最终获得最优策略。用集合 S 、 A 和 R 分别表示状态空间、动作空间和所有智能体的奖励集合。在 t 时刻, 智能体根据其所处状态 $s_t \in S$, 基于决策策略选择处于状态 s_t 时智能体对应的动作 $a_t \in A$, 与环境交互后得到反馈, 即奖励值 r_t , 并进入下一状态 s_{t+1} 。RL 的目标就是通过不断地“试错”得到最优选择策略 $\pi^*(s)$, 表示为 $\pi^*(s) = \operatorname{argmax} Q(s_t, a_t)$ 。值函数 $Q(s_t, a_t)$ 表示智能体在状态 s_t 下执行动作 a_t 后, 获得即时奖励和下一状态的值函数的折扣和, $Q(s_t, a_t)$ 定义为

$$Q(s_t, a_t) = E[r_t + \gamma_t \max_{a_{t+1}} Q(s_{t+1}, a_{t+1})] \quad (19)$$

其中, r_t 表示即时奖励, $\gamma \in [0, 1]$ 是折扣因子, 表示未来奖励值对累积奖励值的影响。

传统的基于值函数的 RL 方法是使用期望值对累积奖励进行建模, 表示为动作价值函数 $Q(s, a)$ 。然而在建模过程中, 可能会丢失一些价值信息, 难以有效适应动态变化的 SAGIN 环境和网络选择过程。因此, 采用基于值分布的 DRL 方法对累积奖励的价值分布进行建模^[25], 以适应动态变化的 SAGIN 选择环境。将累积奖励视为一个随机变量 $Z(s, a)$, 并通过一个离散的概率分布来表示其可能的取值, 从而更全面地捕获环境的复杂性和不确定性。 $Z(s, a)$ 可表示为

$$Z(s, a) = P(v|s, a) \quad (20)$$

其中, v 表示可能的价值, $P(v|s, a)$ 表示在给定状态 s 和动作 a 下, 价值 v 出现的概率。

对于给定的状态-动作对 (s, a) , 其对应的动作价值函数 $Q(s, a)$ 可表示为随机变量 $Z(s, a)$ 的期望值。

$$Q(s, a) = E[Z(s, a)] \quad (21)$$

假设价值分布的范围为 $[v_{\min}, v_{\max}]$, 并均分为 H 个离散支点, 这些点表示 Z 值的可能取值, v_{\min} 和 v_{\max} 分别表示价值分布的最小值和最大值, 以此模拟累积奖励的近似分布。每个等分支集可表示为

$$z_i = v_{\min} + i\Delta z \quad (22)$$

其中, $\Delta z = \frac{v_{\max} - v_{\min}}{H - 1}$, $0 \leq i \leq H - 1$ 。

通过主网络输出当前状态-动作对的值分布, 表示为

$$Z(s_t, a_t; \theta_{\text{main}}) = \{(z_i, p_i(s_t, a_t; \theta_{\text{main}}))\}_{i=0}^{H-1} \quad (23)$$

其中, $p_i(s_t, a_t; \theta_{\text{main}})$ 表示状态-动作对 (s_t, a_t) 的值为 z_i 的概率, θ_{main} 表示主网络参数。

使用主网络选择下一状态 s_{t+1} 的最优动作 a^* , 表示为

$$a^* = \arg \max_{a_{t+1}} \sum_{i=0}^{H-1} z_i p_i(s_{t+1}, a_{t+1}; \theta_{\text{main}}) \quad (24)$$

通过目标网络输出下一状态 s_{t+1} 的值分布, 表示为

$$Z(s_{t+1}, a^*; \theta_{\text{target}}) = \{(z_j, p_j(s_{t+1}, a^*; \theta_{\text{target}}))\}_{j=0}^{H-1} \quad (25)$$

其中, θ_{target} 表示目标网络参数, $0 \leq j \leq H-1$ 。

根据 Bellman 方程, 目标分布可表示为

$$\mathcal{T}z_j = r_t + \gamma_i z_j \quad (26)$$

目标分布的取值范围通常不在 $[v_{\min}, v_{\max}]$ 内, 因此需将更新后的目标分布投影到离散化的支集上, 对目标分布进行截断和投影。

$$\hat{\mathcal{T}}z_j = \max(v_{\min}, \min(v_{\max}, \mathcal{T}z_j)) \quad (27)$$

将目标分布投影在离散支集上, 表示为

$$b_j = \frac{\hat{\mathcal{T}}z_j - v_{\min}}{\Delta z} \quad (28)$$

其中, b_j 表示投影后的离散位置, $l_j = \lfloor b_j \rfloor$ 和 $u_j = \lceil b_j \rceil$ 分别表示目标 $\mathcal{T}z_j$ 投影在离散支集上的下界和上界支点的位置。

然后通过式(29)进行投影更新。

$$m_j = p_j(s_{t+1}, a^*; \theta_{\text{target}})(u_j - b_j) + p_j(s_{t+1}, a^*; \theta_{\text{target}})(b_j - l_j) \quad (29)$$

通过最小化当前分布与目标分布之间的交叉熵损失更新网络参数, 则损失函数可表示为

$$L(\theta_{\text{main}}) = - \sum_{i=0}^{H-1} m_i \log p_i(s_t, a_t; \theta_{\text{main}}) \quad (30)$$

每隔 F 步将主网络参数 θ_{main} 复制给目标网络参数 θ_{target} , 并更新目标网络参数 $\theta_{\text{target}} = \theta_{\text{main}}$ 。

2.3.2 基于 D-DQN 的网络选择

图 4 为基于 D-DQN 的 SAGIN 选择算法原理框架。在此框架下, 通过构建合理的状态空间、动作空间和奖励函数, 智能体能够在与环境的不断交互中学习最优的网络选择策略。

D-DQN 模型由一个输入层、2 个全连接隐藏层和一个输出层构成。输入层接收状态空间的特征表示, 2 个全连接隐藏层分别包含 128 和 64 个神经元, 均采用 ReLU 激活函数进行非线性变换。输出层为全连接层, 其输出维度为动作空间大小与离散化支点数量的乘积, 表示每个动作的概率分布。

将 SAGIN 选择问题建模为马尔可夫决策过程, 包括 4 个关键要素: 智能体、状态空间、动作空间和奖励函数。具体描述如下。

1) 智能体。在本文中, 智能体是指 SAGIN 切换系统中的 MT。

2) 状态空间。将 MT 当前的业务类型 $S_{\text{type}} \in \{0, 1, 2, 3\}$ 、当前选择的网络节点 n 及表示网络状态的 QoS 参数 $k_n \in \{B_n, D_n, J_n, \text{BER}_n\}$ 作为网络选择决策因素, 则状态空间 S 可表示为

$$S = \{S_{\text{type}}, D_n, k_n | n = 1, 2, \dots, N\} \quad (31)$$

3) 动作空间。智能体的动作为选择一个网络节点进行切换。由于 SAGIN 的异构性, 同一时刻可能存在多个网络节点可供 MT 进行切换选择, 但在任何 t 时刻 MT 只能选择一个网络节点进行接入。则动作空间 A 可表示为

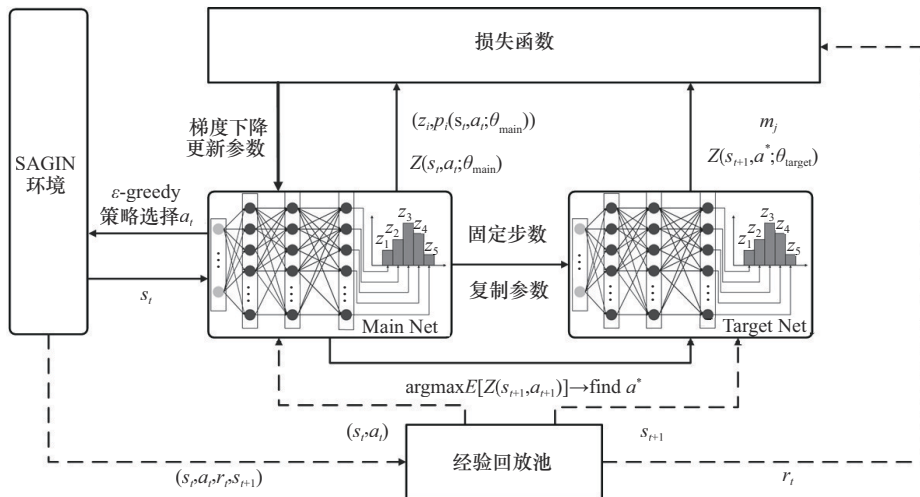


图 4 基于 D-DQN 的 SAGIN 选择算法原理框架

$$A = \{a_i | a_i \in (\mathcal{D}_1, \mathcal{D}_2 \cdots \mathcal{D}_n \cdots \mathcal{D}_N)\} \quad (32)$$

4) 奖励函数。用 r_t 表示智能体在 t 时刻执行动作 a_t 后的奖励。则奖励函数可表示为

$$r_t = U_{n,k,t} - C_{n,m,t} + L_{n,t} \quad (33)$$

基于 D-DQN 的网络选择算法如算法 2 所示。

算法 2 基于 D-DQN 的网络选择算法

输入 学习率 α 、折扣因子 γ 、初始探索率 ε 、迭代次数 X 、批量大小 batch size、目标网络权重参数更新频率 F 、离散支点个数 H

输出 最优网络选择策略

- 1) 初始化神经网络参数 θ_{main} 、 θ_{target} 和经验回放池 Y
- 2) for episode=1,2,⋯,X do
- 3) 初始化状态 s_0
- 4) for $t=1:T$ do
- 5) 利用 ε -greedy 策略选择动作 a_t
- 6) 执行动作 a_t , 获得下一状态 s_{t+1} 和奖励 r_t
- 7) 将 (s_t, a_t, r_t, s_{t+1}) 存储到经验回放池 Y 中
- 8) 从 Y 中随机抽取小批量的经验样本
- 9) 计算目标分布 $\mathcal{T}z$, 将其投影到 H 个离散支点上并计算 m_t
- 10) 根据式(30)计算损失, 并通过反向传播更新主网络参数 θ_{main}
- 11) 每隔 F 步更新目标网络参数 $\theta_{\text{target}} = \theta_{\text{main}}$
- 12) end for
- 13) end for

3 算法仿真与分析

3.1 仿真环境及参数

本文采用卫星工具包 (STK, satellite tool kit) 和 Python 仿真软件构建 SAGIN 仿真场景, 验证本文算法的有效性。SAGIN 仿真场景如图 5 所示, 包括 10 个 LEO 卫星网络节点、6 个 UAV 网络节点、100 个 5G 网络节点以及 100 个具有一定移动速度的 MT。LEO 卫星网络节点选择的是 Starlink 星座的第一阶段一层卫星网络, 它由 72 个轨道组成, 每个轨道上部署有 22 颗卫星, 卫星运行在距地面 550 km 的轨道上, 轨道倾角为 53° 。MT 业务随机产生, 4 种业务类型产生的概率分别为会话类 40%、交互类 30%、流媒体类 20% 和后台类 10%^[19,21]。网络选择决策时间间隔设置为 10 s, 其

他仿真参数如表 3 所示。3 种网络的网络属性参数值如表 4 所示, 4 种业务类型的 QoS 属性参数需求如表 5 所示。

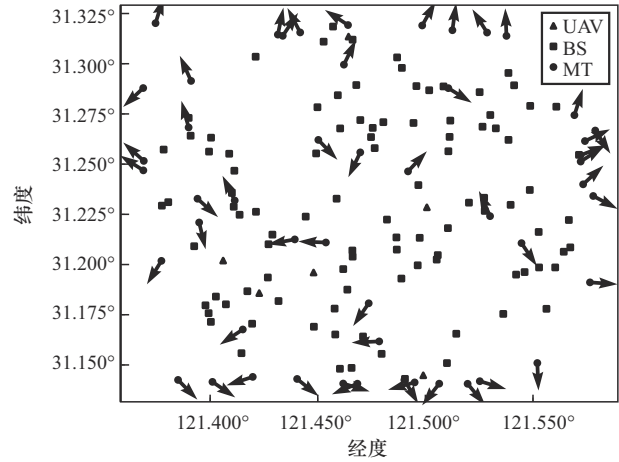


图 5 SAGIN 仿真场景

表 3 仿真参数

	参数	参考数值
环境参数	LEO 卫星网络工作频率/GHz	11.7
	UAV 网络工作频率/GHz	2.4
	5G 网络工作频率/GHz	3.5
	LEO 卫星网络总带宽/MHz	250
	UAV 网络总带宽/MHz	20
	5G 网络总带宽/MHz	50
	LEO 卫星网络节点发射功率/dBm	37
	UAV 网络节点发射功率/dBm	50
	5G 网络节点发射功率/dBm	46
	LEO 卫星网络节点的发射机天线增益/dB	30
	UAV 网络节点的发射机天线增益/dB	20
	5G 网络节点的发射机天线增益/dB	10
	MT 接收天线增益/dB	1
	5G 网络的小区覆盖半径/m	150
	UAV 网络节点高度/m	100
最小通信仰角	10°	
系统等效温度 T_e /K	300	
算法 2 参数	离散支点个数 H /个	51
	网络参数更新频率 F	100
	迭代次数 X /次	1 000
	经验回放池容量	500 000
	批量大小 batch size	64

表 4 网络属性参数值

网络类型	可用带宽/(Mbit·s ⁻¹)	时延/ms	抖动/ms	误码率
LEO 卫星	100	10~100	10~50	20×10 ⁻⁶ ~50×10 ⁻⁶
UAV	20	10~50	10~30	10×10 ⁻⁶ ~100×10 ⁻⁶
5G	50	5~20	1~20	1×10 ⁻⁶ ~10×10 ⁻⁶

表 5 4 种业务类型的 QoS 属性参数需求

QoS 需求	会话类	交互类	流媒体类	后台类
最小带宽 B_{\min} /(Mbit·s ⁻¹)	0.1	0.1	0.3	0.1
最大带宽 B_{\max} /(Mbit·s ⁻¹)	8	6	12	5
最小时延 D_{\min} /ms	10	20	15	40
最大时延 D_{\max} /ms	25	40	30	80
最小抖动 J_{\min} /ms	5	15	10	25
最大抖动 J_{\max} /ms	15	30	25	50
最小误码率 BER _{min}	10×10 ⁻⁶	5×10 ⁻⁶	8×10 ⁻⁶	2×10 ⁻⁶
最大误码率 BER _{max}	35×10 ⁻⁶	25×10 ⁻⁶	25×10 ⁻⁶	12×10 ⁻⁶

3.2 仿真结果

3.2.1 MT 位置预测性能

MT 与 LEO 卫星网络节点之间的可见性关系是建立卫星通信链路的基本条件, 选择地面目标上海市(东经 121.48°、北纬 31.23°)作为观测位置, 图 6 给出了在观测时间内 MT 可见的卫星数量。从图 6 中可看出, MT 可见卫星数量在 27~78 颗之间波动, 这说明有多个 LEO 卫星网络节点可以为 MT 提供服务, 以满足不同 MT 的需求和环境变化。

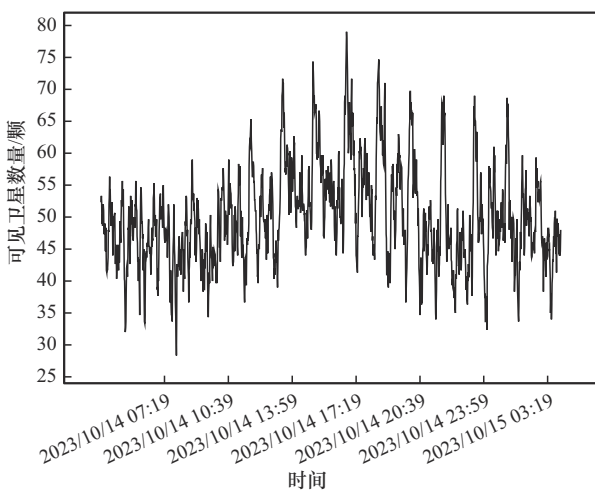


图 6 MT 与卫星网络节点之间的可见性

为评估 SSA-LSTM 算法的 MT 位置预测性能, 本文设计了一个单步预测实验。实验中设置的时间

步长为 10, 即使用连续的 10 个轨迹点作为输入, 预测下一个轨迹点的位置。具体步骤如下。首先, 给定 MT 的前 10 个轨迹点数据, 利用 LSTM 网络模型预测第 11 个轨迹点位置。然后, 以第 2 个轨迹点位置到第 11 个轨迹点位置为新输入, 预测第 12 个轨迹点位置, 以此类推, 进行更长时间的预测。测试数据集采用德国科隆 Koln Trace 车辆轨迹数据集^[26]。其中, 每条轨迹数据包含时间(间隔 1 s)、ID、位置坐标(X,Y)和速度等信息。本文使用了 2 054 条轨迹数据, 共包含 1 529 294 个轨迹点, 并按照 7:1.5:1.5 的比例划分为训练集、验证集和测试集。

本文的 SSA-LSTM 位置预测模型由一个 LSTM 层、一个 Dropout 层和一个全连接隐藏层构成。在前向传播阶段, 该模型首先按序接收输入数据, 其次利用 LSTM 层提取时序特征, 然后通过 Dropout 层设置丢弃率, 以防止过拟合并提升模型的泛化能力。最后, 通过全连接层输出预测值。为了增强模型的稳健性, LSTM 层和全连接隐藏层均采用 L2 正则化。在反向传播阶段, 采用梯度下降法, 使用 Adam 优化器调整权重参数, 以优化预测模型。

SSA 参数设置如下。麻雀种群数量为 20, 发现者占比 20%, 预警者占比 10%, 最大迭代次数为 30。LSTM 网络参数设置如下。学习率搜索范围为 [0.001, 0.01], LSTM 层神经元个数搜索范围为 [1, 200],

全连接隐藏层神经元个数搜索范围为[1,200], 丢弃率搜索范围为[0,0.5], 训练次数 (Epoch) 为 30, 批量大小为 256, L2 正则化因子为 0.000 1。最终确定最优超参数组合为学习率 0.004, LSTM 层的神经元个数 58, 全连接隐藏层神经元个数 149, 丢弃率 0。

图 7 给出了 SSA-LSTM 模型使用最优超参数组合的训练过程。从图 7(a)和图 7(b)可以看出, 随着训练次数 (Epoch) 的增加, 训练集与验证集的损失值逐渐下降, 准确率逐渐升高。当 Epoch 达到 10 次时性能趋于稳定, 稳定状态下的训练集与验证集损

失值分别为 $3.658 3 \times 10^{-5}$ 和 $4.049 8 \times 10^{-5}$, 准确率分别为 0.997 8 和 0.996 6。

基于 SSA-LSTM 模型训练得到的 MT 位置预测性能结果如图 8 所示。从图 8(a)~图 8(c)可以看出, MT 的真实轨迹与预测轨迹较为贴合。图 8(d)为 MT 位置预测误差曲线, 误差均值为 12.873 m, 表明本文算法具有较好的预测性能。

将本文 SSA-LSTM 位置预测模型与反向传播 (BP, back propagation) 模型、LSTM 模型、门控循环单元 (GRU, gate recurrent unit) 模型、双向门控

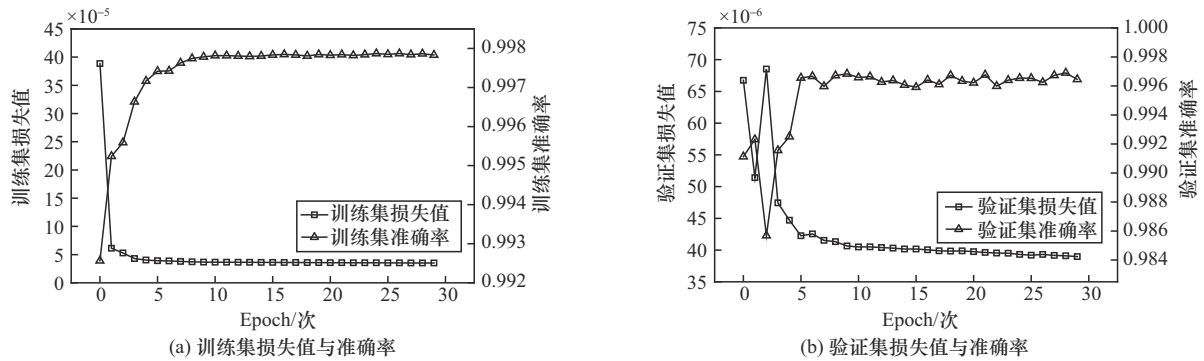


图 7 损失函数和准确率变化曲线

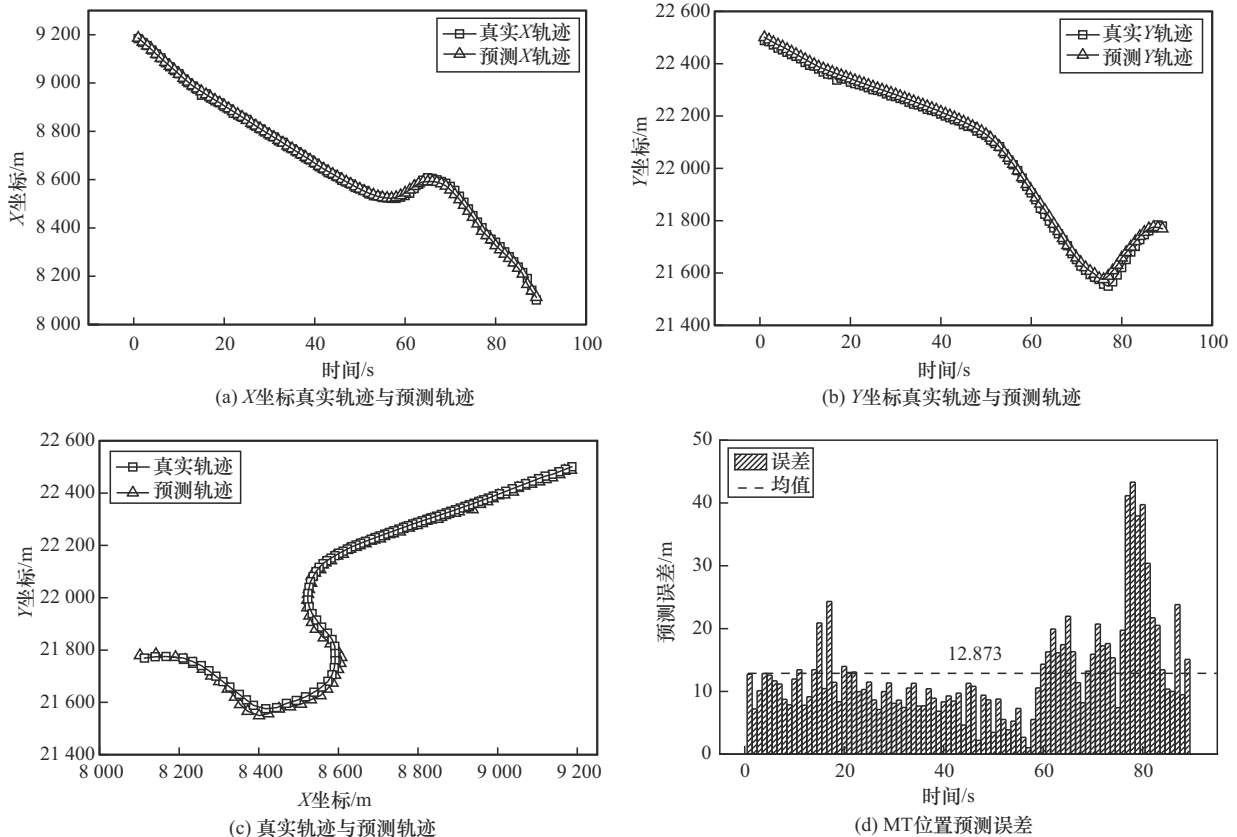


图 8 MT 位置预测性能结果

循环单元 (Bi-GRU, bidirectional gated recurrent unit) 模型和双向长短期记忆 (Bi-LSTM, bidirectional long short-term memory) 网络模型进行对比实验。为了公正地评价各个模型的预测性能, 采用 5 种误差评价指标: 决定系数 (R2, r-squared)、均方误差 (MSE)、均方根误差 (RMSE, root mean square error)、平均绝对误差 (MAE, mean absolute error) 和平均绝对百分比误差 (MAPE, mean absolute percentage error)。鉴于测试数据集包含多条车辆轨迹, 因此需取所有测试结果的平均值作为各模型的性能评价指标。表 6 展示了 6 种不同模型对不同车辆轨迹数据的预测性能。可以看出, 本文模型相较于其他 5 种模型的预测性能表现更好。

表 6 不同模型的预测性能比较

模型	R2	MSE	RMSE	MAE	MAPE
BP	0.965 3	243.230 1	14.546 8	10.368 8	0.090 0%
GRU	0.973 6	192.055 1	11.663 2	9.068 2	0.077 8%
LSTM	0.982 9	137.473 2	10.370 6	7.655 4	0.065 7%
Bi-GRU	0.984 0	154.227 5	11.142 3	8.709 9	0.076 8%
Bi-LSTM	0.977 4	137.904 6	10.947 0	7.995 8	0.067 9%
SSA-LSTM	0.990 2	122.970 9	9.579 8	6.952 8	0.059 4%

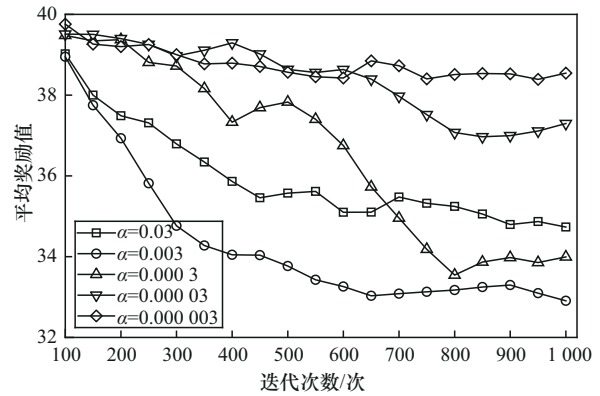
3.2.2 网络切换性能

本节将本文所提基于 SSA-LSTM 和 D-DQN 的网络切换算法 (简称 SL-D-DQN 算法) 与 D3QN 算法^[21]、LB-DDQN 算法^[27]、PER-DDQN 算法^[22] 和 Q-Learning 算法在本文应用场景下进行对比实验。

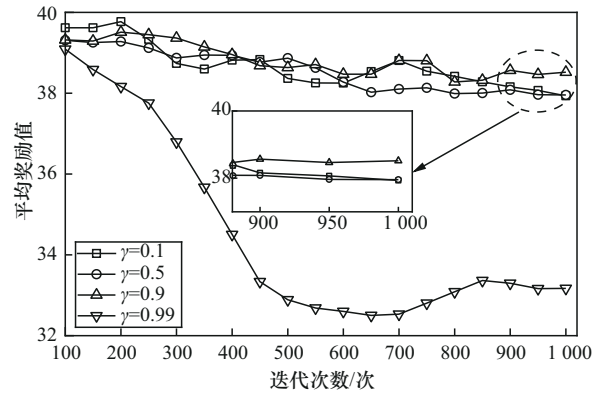
图 9(a)~图 9(c) 分别给出了本文所提 SL-D-DQN 算法在不同学习率 α 、折扣因子 γ 、初始探索率 ϵ 下的平均奖励值。从图 9 中可以观察到, 在学习率 $\alpha=0.000\ 003$ 、折扣因子 $\gamma=0.9$ 和初始探索率 $\epsilon=0.5$ 时的平均奖励值最高。这说明合理地选择超参数组合可以显著地提升算法性能。

图 10 为不同算法的平均奖励值对比。从图 10 中可以看出, SL-D-DQN 算法获得的平均奖励值高于 D3QN 和 Q-Learning 算法。这是因为 SL-D-DQN 算法输出价值分布, 考虑了整个价值分布而不仅仅是平均值, 能够更好地建模环境的不确定性, 选择长期累积奖励值最大的网络进行接入。SL-D-DQN 算法的平均奖励值相比 LB-DDQN 和 PER-DDQN 算法较低, 这是因为 SL-D-DQN 算法在奖励

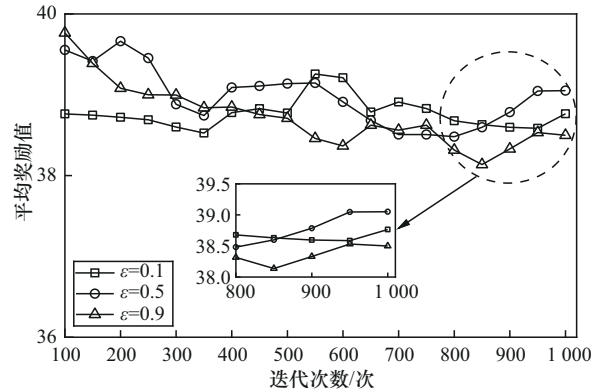
函数中额外考虑了切换代价, 并将其作为一个惩罚因素在奖励中减去。



(a) 不同学习率 α 下的平均奖励值



(b) 不同折扣因子 γ 下的平均奖励值



(c) 不同初始探索率 ϵ 下的平均奖励值

图 9 不同 α, γ, ϵ 值对平均奖励值的影响

图 11 给出了 4 种业务类型下不同算法的 MT 平均接入阻塞概率对比。从图 11 中可以观察到, 当业务到达率 $\lambda < 0.3$ 时, 所有业务类型的 MT 平均接入阻塞概率都趋近于 0, 网络几乎没有出现阻塞的情况。这是因为较低的 λ 值产生的业务请求较少, 网络资源充足, 能够满足当前业务需求。然而, 随着 λ 值的增大, 4 种业务的 MT 平均接入阻塞概

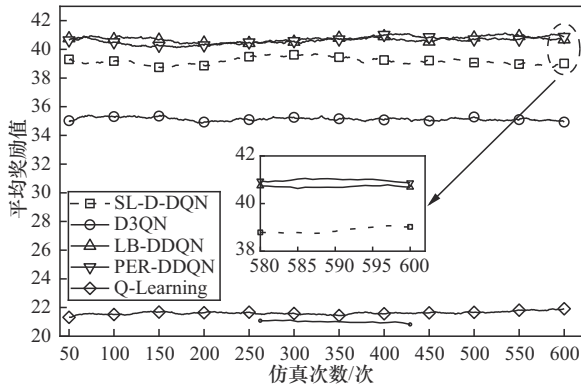
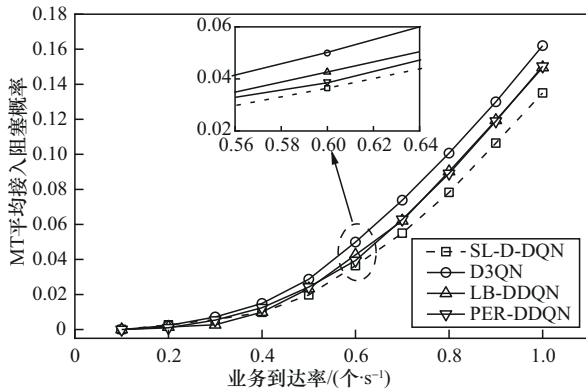


图 10 不同算法的平均奖励值对比

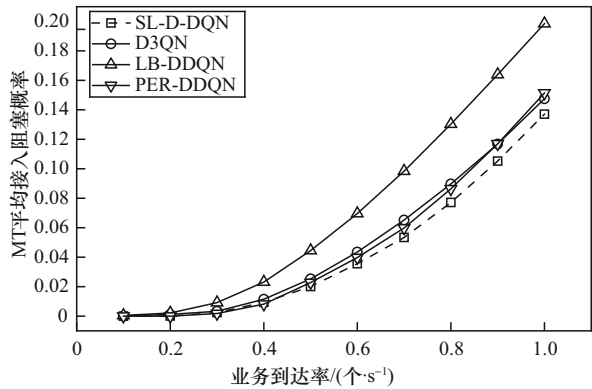
率曲线均呈递增趋势。相比于其他算法,在4种业务类型下,SL-D-DQN算法的MT平均接入阻塞概率均为最低。这是因为SL-D-DQN算法综合考虑了MT的业务类型、不同业务的QoS需求和网络状态来保障MT的QoS需求,并以最大化综合效用函数值为切换算法设计目标。当 $\lambda=1$ 时,会话类业务下SL-D-DQN算法的MT平均接入阻塞概率比D3QN、LB-DDQN和PER-DDQN算法分别降低了16.66%、9.70%和10.14%;交互类业务下分

别降低了7.10%、30.96%和9.46%;流媒体类业务下分别降低了22.03%、9.73%和26.28%;后台类业务下分别降低了13.62%、38.64%和25.05%。

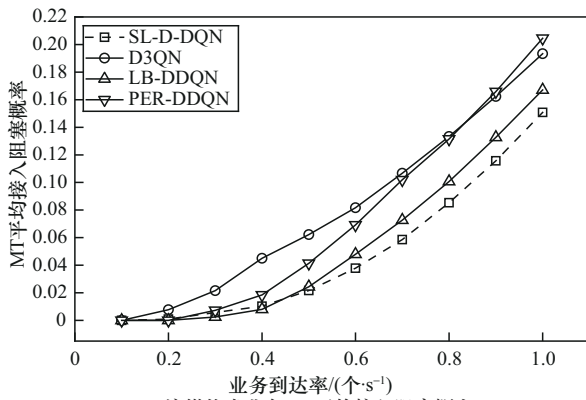
图12为业务到达率 λ 与系统平均吞吐量之间的关系。从图12中可以看出,当 λ 值较小时,5种算法的吞吐量增长速度较快。这是因为 λ 值较小情况下的网络带宽资源相对充裕,MT不易被阻塞。随着 λ 值的增大,可用的网络带宽资源越来越有限,各算法的吞吐量增长速度趋于缓慢。当 $\lambda>0.4$ 时,SL-D-DQN、LB-DDQN和PER-DDQN算法具有较高的吞吐量。当 $\lambda=1$ 时,SL-D-DQN、D3QN、LB-DDQN、PER-DDQN和Q-Learning算法的平均吞吐量分别为3 521 Mbit/s、3 170 Mbit/s、3 352 Mbit/s、3 382 Mbit/s和2 358 Mbit/s,SL-D-DQN算法的平均吞吐量比D3QN、LB-DDQN、PER-DDQN和Q-Learning算法分别提高了11.07%、5.04%、4.11%和49.26%,表明SL-D-DQN算法在高业务量下的吞吐量性能更优。这是因为本文算法在网络选择时考虑了MT业务需求、候选网络QoS属性、切换代价和网络负载多方面因素,较好地实现了候选网络



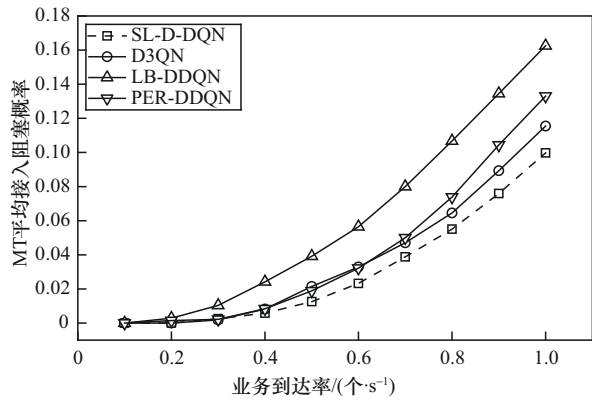
(a) 会话类业务MT平均接入阻塞概率



(b) 交互类业务MT平均接入阻塞概率



(c) 流媒体类业务MT平均接入阻塞概率



(d) 后台类业务MT平均接入阻塞概率

图 11 不同算法的MT平均接入阻塞概率对比

的负载均衡，从而提升了系统吞吐量性能。LB-DDQN 和 PER-DDQN 算法平均吞吐量高于 D3QN 算法，这是因为这 2 种算法将负载系数作为切换决策因素之一，实现了切换决策和负载均衡。

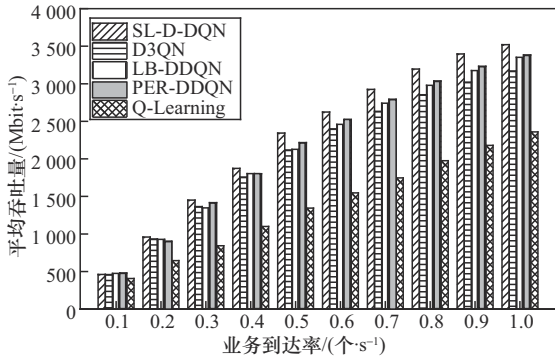


图 12 系统平均吞吐量

图 13 为不同算法的平均切换次数对比。从图 13 中可以看出，随着网络切换决策次数的增加，MT 平均切换次数也随之增加。当决策次数为 600 次时，5 种算法的平均切换次数由高到低分别是 878、785、701、612 和 412 次。与其他算法相比，SL-D-DQN 算法的平均切换次数最少。这是因为 SL-D-DQN 算法首先采用位置预测模型对 MT 下一时刻的位置进行预测，并基于此设计了有效的切换触发机制。同时针对不同业务对网络属性参数的差异化需求，为其设计了个性化的效用函数，有效地满足了各种业务类型的 MT 的 QoS 需求，避免了不必要的切换。此外，在网络选择时考虑了切换代价，进一步减少了切换次数。

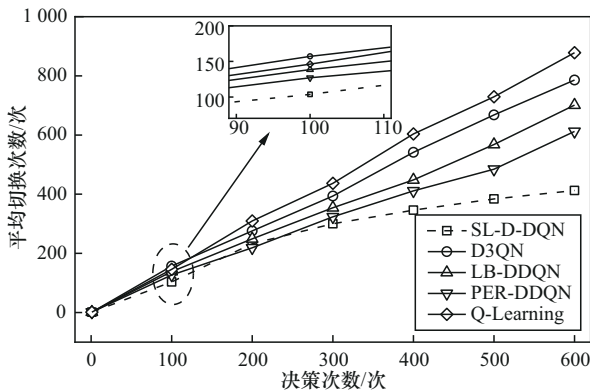


图 13 不同算法的平均切换次数对比

图 14 为不同算法的平均切换失败概率对比。从图 14 中可以看出，随着 MT 业务到达率的增加，平

均切换失败概率也逐渐增大。这是因为当 MT 业务较少时，网络资源比较充足，MT 之间竞争性较低，更易实现网络切换。SL-D-DQN 算法具有较低的切换失败概率，因为该算法通过对 MT 位置的精确预测能够预先估计 MT 的最佳切换时机。此外，基于奖励函数执行切换决策时，综合考虑了 QoS 属性参数、切换代价和负载等多个影响切换的效用因素，为 MT 提供了最佳的切换决策。D3QN 算法的平均切换失败概率较高，这是因为其未考虑候选网络的负载情况，导致候选网络节点因为接入过载而拒绝接收新的 MT 切换请求，增加了切换失败的概率。而对于 Q-Learning 算法，随着 MT 数量的增加，由于其受 Q 值表存储能力限制，难以高效地处理所有状态和相应的动作，导致切换失败概率增加。

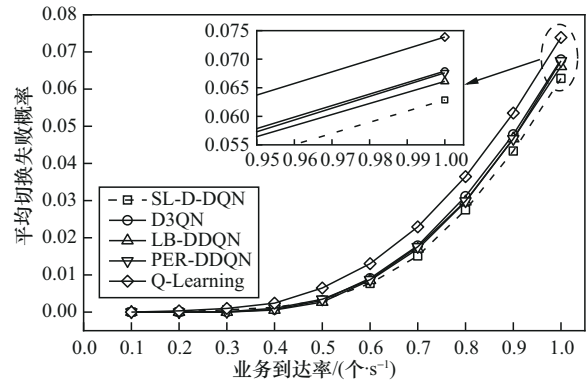


图 14 不同算法的平均切换失败概率对比

图 15 为不同算法的平均切换时延对比。从图 15 中可以看出，Q-Learning 算法的平均切换时延最高，DQN 类算法在降低切换时延方面表现出色，SL-D-DQN 算法的平均切换时延最低。这是因为随着状态空间的增大，Q-Learning 算法的计算复杂度会相应增加。DQN 类算法使用神经网络来逼近 Q 函数，避免了直接查询 Q 值表所耗费的时间，从而降低了切换时延。SL-D-DQN 算法通过预测 MT 周围候选网络节点的链路质量提前进行切换触发，这种预先触发机制显著减少了从一个网络切换到另一个网络所需的时间，有效降低了切换时延。

3.2.3 复杂度分析

由 2.2 节和 2.3 节可知，SL-D-DQN 算法主要由两部分组成。第一部分为使用 LSTM 网络进行位置预测。通常情况下，LSTM 网络的时间复杂度由序列长度 T_{step} 、输入数据维度 N_{LSTM} 和隐藏层神经元数量 M_{LSTM} 组成，即时间复杂度为 $O(T_{step}(N_{LSTM}M_{LSTM} +$

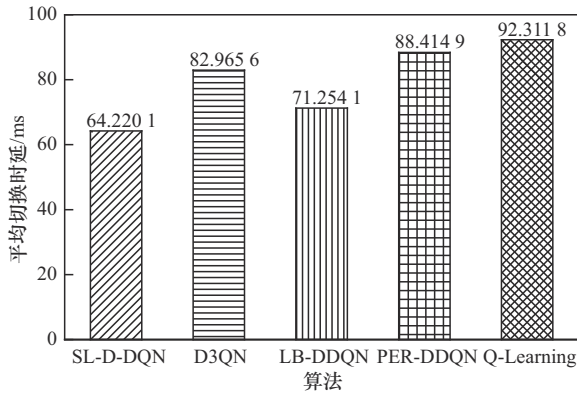


图 15 不同算法的平均切换时延对比

M_{LSTM}^2)。第二部分为使用 D-DQN 算法进行网络选择, DQN 类算法的时间复杂度主要来源于神经网络的训练过程,即前向传播和反向传播。根据本文中 DQN 的网络结构,输入层维度记为 N_s , 2 个隐藏层神经元数量分别为 h_1 和 h_2 , 输出层对应的动作空间大小为 N , 经验回放池大小为 Y_m , batch size 大小为 U , 离散支点的个数为 H 。因此, D-DQN 算法的时间复杂度为 $O(U(2(N_s h_1 + h_1 h_2 + h_2 N H) + U N H))$, 则 SL-D-DQN 算法的时间复杂度为 $O(T_{step}(N_{LSTM} M_{LSTM} + M_{LSTM}^2) + O(U(2(N_s h_1 + h_1 h_2 + h_2 N H) + U N H))$ 。LB-DDQN 算法使用 2 个 Q 网络, 因此其时间复杂度为 $O(U(2(N_s h_1 + h_1 h_2 + h_2 N)))$ 。PER-DDQN 算法引入了优先级采样机制, 从而在训练过程中增加了经验采样的复杂度, 则其时间复杂度为 $O(U(2(N_s h_1 + h_1 h_2 + h_2 N))) + O(B \log Y_m)$ 。D3QN 算法的主要区别在于 Q 值的计算方式, Q 值被分解为状态价值函数和优势价值函数。因此, D3QN 算法的时间复杂度为 $O(U(2(N_s h_1 + h_1 h_2 + h_2 h_v + h_2 h_a N)))$, 其中, h_v 和 h_a 分别表示状态价值和优势价值分支的神经元个数。尽管 SL-D-DQN 算法的计算复杂度较高, 但它在网络选择过程中能够提供更加细致的 Q 值分布估计, 避免了不必要的频繁切换, 有效减少了切换次数和切换时延。

4 结束语

本文所提基于 SL-D-DQN 的网络切换算法, 适用于不同业务类型的 MT 在 SAGIN 中的网络切换问题。该算法将网络切换决策过程分为 2 个阶段, 在切换触发阶段, 通过 SSA-LSTM 学习 MT 的历史移动轨迹构造位置预测模型, 然后根据该模型预测 MT 下一时刻的位置并估计链路质量, 以此设计切

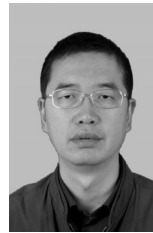
换触发机制, 及时触发切换。当切换触发时, 采用基于 D-DQN 的网络选择算法选择目标切换网络节点。首先综合考虑 MT 的业务类型、网络切换代价和网络负载均衡, 并基于效用函数理论设计了综合的奖励函数, 在保证 MT 通信 QoS 的同时减少不必要的切换。仿真结果进一步验证了本文算法的有效性。在未来工作中, 将进一步细分 SAGIN 的应用场景 (如具有快时变特性的高速铁路场景), 结合不同场景的网络需求, 设计更高可靠, 更低时延的低复杂度网络切换算法。

参考文献:

- [1] BARAKABITZE A A, WALSH R. SDN and NFV for QoE-driven multimedia services delivery: the road towards 6G and beyond networks[J]. Computer Networks, 2022, 214: 109133.
- [2] FU I K, CHARBIT G, MEDLES A, et al. Satellite and terrestrial network convergence on the way toward 6G[J]. IEEE Wireless Communications, 2023, 30(1): 6-8.
- [3] WANG C X, YOU X H, GAO X Q, et al. On the road to 6G: visions, requirements, key technologies, and testbeds[J]. IEEE Communications Surveys & Tutorials, 2023, 25(2): 905-974.
- [4] RAY P P. A review on 6G for space-air-ground integrated network: key enablers, open challenges, and future direction[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(9): 6949-6976.
- [5] 张海君, 陈安琪, 李亚博, 等. 6G 移动网络关键技术[J]. 通信学报, 2022, 43(7): 189-202.
- [6] ZHANG H J, CHEN A Q, LI Y B, et al. Key technologies of 6G mobile network[J]. Journal on Communications, 2022, 43(7): 189-202.
- [7] ZHU X M, JIANG C X. Integrated satellite-terrestrial networks toward 6G: architectures, applications, and challenges[J]. IEEE Internet of Things Journal, 2022, 9(1): 437-461.
- [8] SHENG J, CAI X Q, LI Q Y, et al. Space-air-ground integrated network development and applications in high-speed railways: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 10066-10085.
- [9] 钱志鸿, 冯一诺, 孙佳妮, 等. 基于 DA 优化模糊神经网络的异构无线网络接入选择算法[J]. 通信学报, 2020, 41(12): 118-127.
- [10] QIAN Z H, FENG Y N, SUN J N, et al. Access selection algorithm for heterogeneous wireless network based on DA optimized fuzzy neural network[J]. Journal on Communications, 2020, 41(12): 118-127.
- [11] GUO C, GONG C, XU H T, et al. A dynamic handover software-defined transmission control scheme in space-air-ground integrated networks[J]. IEEE Transactions on Wireless Communications, 2022, 21(8): 6110-6124.
- [12] 胡欣, 宋航宇, 刘帅军, 等. 基于时间演进图的 LEO 星间切换实时预测及更新方法[J]. 通信学报, 2018, 39(10): 43-51.
- [13] HU X, SONG H Y, LIU S J, et al. Real-time prediction and updating method for LEO satellite handover based on time evolving graph[J]. Journal on Communications, 2018, 39(10): 43-51.

- [11] ZHOU Y T, MENG R, XIONG H C, et al. Evolutionary game-based vertical handover strategy for space-air-ground integrated network[C]// Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE Press, 2023: 1-6.
- [12] FAN K X, FENG B W, ZHANG X L, et al. Network selection based on evolutionary game and deep reinforcement learning in space-air-ground integrated network[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(3): 1802-1812.
- [13] 杨喆, 邓立宝, 狄原竹, 等. 超密集异构无线网络中基于移动轨迹预测的网络切换算法[J]. 电子与信息学报, 2023, 45(12): 4280-4291.
YANG Z, DENG L B, DI Y Z, et al. Network switching algorithm based on mobile trajectory prediction in ultra-dense heterogeneous wireless networks[J]. Journal of Electronics & Information Technology, 2023, 45(12): 4280-4291.
- [14] 马彬, 王双双, 陈海波. 基于区间二型模糊神经网络的垂直切换算法[J]. 电子学报, 2021, 49(5): 928-935.
MA B, WANG S S, CHEN H B. Vertical handover algorithm based on interval type-2 fuzzy neural network[J]. Acta Electronica Sinica, 2021, 49(5): 928-935.
- [15] XU H H, LI D S, LIU M L, et al. QoE-driven intelligent handover for user-centric mobile satellite networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(9): 10127-10139.
- [16] WANG F, JIANG D D, WANG Z H, et al. Seamless handover in LEO based non-terrestrial networks: service continuity and optimization[J]. IEEE Transactions on Communications, 2023, 71(2): 1008-1023.
- [17] ZHU A Q, GUO S T, LIU B, et al. Adaptive multiservice heterogeneous network selection scheme in mobile edge computing[J]. IEEE Internet of Things Journal, 2019, 6(4): 6862-6875.
- [18] SUN J N, QIAN Z H, WANG X, et al. ES-DQN-based vertical handoff algorithm for heterogeneous wireless networks[J]. IEEE Wireless Communications Letters, 2020, 9(8): 1327-1330.
- [19] TAVANA M, SOROOSHIAN S, REZAEI H, et al. A novel fuzzy scenario-based stochastic general best-worst method[J]. Expert Systems with Applications, 2024, 252: 124246.
- [20] HONARVAR R, ZOLGHADRASLI A, MONEMI M. Context-oriented performance evaluation of network selection algorithms in 5G heterogeneous networks[J]. Journal of Network and Computer Applications, 2022, 202: 103358.
- [21] XIE J L, ZHU B H, LI C R. Research of 5G HUDN network selection algorithm based on dueling-DDQN[J]. EURASIP Journal on Wireless Communications and Networking, 2023, 2023(1): 113.
- [22] 唐宏, 刘小洁, 甘陈敏, 等. 超密集网络中基于改进DQN的接入选择算法[J]. 哈尔滨工业大学学报, 2023, 55(5): 107-113.
TANG H, LIU X J, GAN C M, et al. Access selection algorithm based on improved DQN for ultra-dense networks[J]. Journal of Harbin Institute of Technology, 2023, 55(5): 107-113.
- [23] XUE J K, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.
- [24] XU J P, AI B. Deep reinforcement learning for handover-aware MPTCP congestion control in space-ground integrated network of railways[J]. IEEE Wireless Communications, 2021, 28(6): 200-207.
- [25] LI R X, SHANG Z W, ZHENG C H, et al. Efficient distributional reinforcement learning with Kullback-Leibler divergence regularization[J]. Applied Intelligence, 2023, 53(21): 24847-24863.
- [26] UPPOOR S, TRULLOLS-CRUCES O, FIORE M, et al. Generation and analysis of a large-scale urban vehicular mobility dataset[J]. IEEE Transactions on Mobile Computing, 2014, 13(5): 1061-1075.
- [27] WU D F, HUANG C H, YIN Y B, et al. LB-DDQN for handover decision in satellite-terrestrial integrated networks[J]. Wireless Communications and Mobile Computing, 2021, 2021(1): 5871114.

[作者简介]



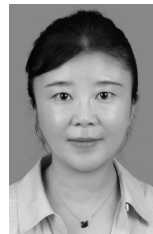
谢健骊 (1972-), 男, 甘肃陇西人, 博士, 兰州交通大学教授、博士生导师, 主要研究方向为高铁智能通信、认知无线电、铁路物联网技术等。



陈龙 (2001-), 男, 甘肃天水人, 兰州交通大学硕士生, 主要研究方向为空地一体化网络移动性管理、网络接入选择。



张泽鹏 (1998-), 男, 甘肃白银人, 兰州交通大学博士生, 主要研究方向为智能超表面、轨道交通通信、智能隐蔽通信。



李翠然 (1975-), 女, 山西黎城人, 博士, 兰州交通大学教授、博士生导师, 主要研究方向为高速铁路智能无线通信、无线传感器网络、协同通信技术等。