

基于强化学习的服务功能链部署综述

爱因坦¹ 阮宏伟¹ 刘翰林¹ 李华^{1,2,3}

¹(内蒙古大学计算机学院 呼和浩特 010021)

²(生态大数据教育部工程研究中心(内蒙古大学) 呼和浩特 010021)

³(内蒙古自治区云计算与服务软件工程实验室(内蒙古大学) 呼和浩特 010021)

(aiyintan@mail.imu.edu.cn)

Survey on Reinforcement Learning-Based Service Function Chain Deployment

Ai Yintan¹, Ruan Hongwei¹, Liu Hanlin¹, and Li Hua^{1,2,3}

¹(College of Computer Science, Inner Mongolia University, Hohhot 010021)

²(Engineering Research Center of Ecological Big Data (Inner Mongolia University), Ministry of Education, Hohhot 010021)

³(Inner Mongolia Engineering Laboratory for Cloud Computing and Service Software (Inner Mongolia University), Hohhot 010021)

Abstract With the rapid development of emerging network technologies such as B5G/6G and Internet of things, network services present characteristics of diverse business types, differentiated quality requirements, and dynamic resource demands. Under the software defined networking (SDN) and network function virtualization (NFV) paradigm, service function chaining (SFC) technology enables customized network service deployment through flexible orchestration of virtual network functions (VNF). However, SFC deployment in dynamic network environments faces challenges such as large decision space, and complex and changing environments. Reinforcement learning (RL) demonstrates significant advantages in solving SFC deployment problems due to its capability to adaptively learn complex environmental characteristics and make dynamic decisions. Although research on applying RL to SFC deployment has made certain progress, a systematic and comprehensive review analysis remains lacking. To address this gap, the fundamental concepts and technical architecture of SFC deployment are elaborated, along with a specific introduction to the RL-based SFC deployment framework. Subsequently, from the perspectives of algorithm design, application scenarios, and optimization strategies, the research progress and innovative applications of reinforcement learning in three key phases—SFC placement, scheduling, and reconfiguration—are systematically examined and thoroughly analyzed. Finally, the advantages and limitations of existing research in terms of algorithm design, performance optimization, and practical deployment are summarized, while technical challenges and future development trends in this field are analyzed.

Key words service function chaining (SFC); reinforcement learning; resource optimization; virtual network function; intelligent deployment

摘要 在软件定义网络和网络功能虚拟化范式下,服务功能链(service function chaining, SFC)技术通过

收稿日期: 2025-01-01; 修回日期: 2025-08-11

基金项目: 国家自然科学基金项目(61862047, 62262047); 内蒙古科技计划项目(201802028, 2020GG0186); 支持地方高校改革发展资金(学科建设)、内蒙古自治区一流学科科研专项项目(YLXKZX-ND-036); 内蒙古重点研发和成果转化项目(2025YFHH0101)

This work was supported by the National Natural Science Foundation of China (61862047, 62262047), the Inner Mongolia Science and Technology Program (201802028, 2020GG0186), the Fund of Supporting the Reform and Development of Local Universities (Disciplinary Construction), the Special Research Project of First-class Discipline of Inner Mongolia A. R. of China (YLXKZX-ND-036), and the Inner Mongolia Key Research and Development and Achievement Transformation Project (2025YFHH0101).

通信作者: 李华(cslihua@imu.edu.cn)

灵活编排虚拟网络功能实现了网络服务的定制化部署。然而,动态网络环境下 SFC 部署面临着决策空间大、环境复杂多变等挑战。强化学习因其自适应学习复杂环境特征并动态决策的能力,在解决 SFC 部署问题上展现出显著优势。该研究首先阐述了 SFC 部署的基本概念与技术架构,并具体介绍了基于强化学习的 SFC 部署框架。随后,从算法设计、应用场景和优化策略等角度,梳理并深入分析了强化学习在 SFC 放置、调度和重配置 3 个关键阶段中的研究进展与创新应用。最后,总结了现有研究在算法设计、性能优化和实际部署方面的优势与局限,并分析了该领域的技术挑战与未来发展趋势。

关键词 服务功能链;强化学习;资源优化;虚拟网络功能;智能部署

中图法分类号 TP393

DOI: 10.7544/issn1000-1239.202550002 **CSTR:** 32373.14.issn1000-1239.202550002

随着互联网技术的迅速发展,全球社会正朝着更紧密互联、数字化和多元化的方向迈进。互联网已深度融入人们的日常生活、工作和学习之中,成为推动技术创新和经济增长的关键驱动力。据统计,2024 年已有 55 亿人接入互联网,占全球人口的 68%,相比 2023 年的 65% 又有显著增长^[1]。与此同时,物联网、工业互联网等新型应用场景正驱动网络服务向更加智能化、自动化的方向演进^[2]。在此背景下,传统网络在应对不断增长的网络流量和日益提升的服务质量(quality of service, QoS)需求方面,正面临服务部署效率不足、资源利用率偏低以及管理灵活性有限等多重挑战。

网络功能虚拟化(network function virtualization, NFV)和软件定义网络(software defined networking, SDN)的出现为解决传统网络架构的局限性提供了新思路^[3-4]。NFV 通过将网络功能从专用硬件解耦,实现了网络服务的虚拟化部署^[3];SDN 则通过控制平面和数据平面分离,实现了网络流量的灵活调度^[4]。在此基础上发展起来的服务功能链(service function chaining, SFC)技术,能够将虚拟网络功能(virtual network function, VNF)按序连接,为用户提供定制化的网络服务^[5]。NFV 和 SDN 环境的融合为网络资源的灵活调度和按需分配提供了技术基础,使得基于多维度优化目标(如时延、带宽利用率、能耗等)的智能部署成为可能^[6]。

然而,在复杂多变的网络环境下,SFC 部署面临着诸多挑战:1)决策空间庞大,传统优化方法难以在有限时间内得到满意解;2)网络状态动态变化,静态部署策略难以适应;3)多样化服务需求与有限网络资源之间的矛盾日益突出;4)QoS 保证与资源利用效率的平衡难以把握^[7]。这些挑战促使研究者探索新的智能优化方法。强化学习(reinforcement learning, RL)^[8] 凭借其自适应学习和动态决策能力,在 SFC 部

署中展现出独特优势。特别是随着深度学习技术的突破,基于深度强化学习(deep reinforcement learning, DRL)^[7]、图神经网络(graph neural network, GNN)辅助的强化学习^[9]以及多智能体强化学习(multi-agent reinforcement learning, MARL)^[10]等方法在解决大规模动态优化问题上取得了显著进展。

根据欧洲电信标准协会(European Telecommunications Standards Institute, ETSI)发布的面向下一个 10 年 NFV 演进白皮书的分析,SFC 作为 NFV 的关键应用场景,其发展将从自动化向智能化演进,人工智能将在服务链编排和优化中发挥越来越核心的作用^[11]。特别是在 B5G/6G、边缘计算、工业互联网等新型场景中,SFC 部署需要更加灵活精准地响应多样化的服务需求,这也使得基于强化学习的 SFC 部署方法成为亟需探索的重要方向^[12]。

为系统化地研究基于强化学习的服务功能链部署进展,本文对现有文献进行了全面的调研和深入分析,如表 1 所示。研究发现,早期的相关综述主要聚焦于 SFC 的基本架构、实现方法和优化策略^[6,13-16],对基于强化学习的方法分析较少。近年来,虽有部分文献开始关注强化学习在 SFC 以及虚拟网络功能转发图 VNF-FG(virtual network function forwarding graph)部署中的应用^[12,17-18],但缺乏系统性和全面性的分析。此外,现有研究在 SFC 的放置、调度和重配置等关键环节的优化上关注不足,对算法在不同场景下的适用性及局限性也未进行充分探讨。

本文系统分析并总结了基于强化学习的服务功能链部署研究现状,围绕 SFC 的放置、调度和重配置 3 个关键阶段,梳理了现有研究的主要进展与成果。同时,从算法设计、应用场景和优化策略等角度,深入探讨了强化学习在 SFC 部署中的创新应用,揭示了优化策略的演进规律及不同部署阶段之间的关联性。这些分析不仅为复杂网络环境下的算法选择

Table 1 Comparison of Our Paper and Other Service Function Chaining Relative Surveys

表 1 本文与其他服务功能链相关综述的对比

相关综述	年份	综述范围	基于强化学习的 SFC 分析情况			
			放置	调度	重配置	分析深度
文献 [6]	2016	SFC 定义、架构、用例、优化策略分类	×	×	×	无
文献 [13]	2018	NFV 架构、NFV 环境下 SFC 部署方法	×	×	×	无
文献 [14]	2020	NFV/SDN 架构下的 SFC 定义、架构、优化策略分类	△	×	×	无
文献 [15]	2020	NFV 架构下资源分配、VNF-FG 嵌入问题 (VNF 放置、流量路由)	△	×	×	无
文献 [16]	2021	NFV 架构下资源管理 (组链、部署、调度和按需缩放)	×	△	×	无
文献 [17]	2021	NFV 资源分配问题定义及建模、VNF-FG 嵌入策略分类	△	×	×	浅
文献 [18]	2022	分布式场景下 SFC 放置、优化策略分类	△	×	×	浅
文献 [19]	2023	SFC 用例、安全机制 (架构安全、解决方案、安全管理)	×	×	×	无
文献 [12]	2024	B5G/6G 网络中 VNF-FG 嵌入、优化方法分类	△	△	×	中
本文	2024	基于强化学习的 SFC 部署架构以及 (放置、调度、重配置)	√	√	√	深

注：“√”表示详细分析；“△”表示简要提及；“×”表示未涉及。“无”表示未涉及；“浅”表示仅做简单介绍；“中”表示有专门章节但不系统；“深”表示系统全面分析。

提供了理论指导,也为实际应用中的技术实现提供了参考支持。此外,系统分析了现有研究的优势与局限性,总结了当前研究中面临的关键挑战,并提出了未来可能的研究方向。

1 服务功能链概述

服务功能链是一种先进且灵活的网络技术,旨在根据具体的服务需求与网络资源的可用性,对数据流进行分类管理并动态部署策略,从而提供定制化的网络服务^[5]。SFC 通过将网络流量引导至一系列按特定顺序排列的网络功能模块,确保数据流逐步经过处理以实现既定的服务效果^[6]。

本节首先介绍 NFV 的基本概念,在此基础上阐述 SFC 技术的基本架构,并分析 SFC 部署的关键流程与步骤,最后讨论强化学习在 SFC 部署中的应用框架。

1.1 NFV 概述

NFV 是一种通过解耦网络功能与专用硬件,从而提升网络灵活性与扩展性的关键技术^[3]。其核心理念是将传统依赖专用硬件实现的网络功能抽象并虚拟化,使其能够以 VNF 的形式运行于标准化、高性能的通用服务器之上,从而打破传统网络设备与软件的捆绑。通过这一解耦过程,NFV 能够支持计算、存储和网络资源的按需动态分配,实现资源的高效共享与优化利用,进而提升整体硬件资源的利用率。此外,NFV 架构的引入不仅加速了新型网络服务的开发与部署,提升了网络管理的敏捷性与运维效率,同时还能有效降低资本支出 (capital expenditure,

CAPEX) 和运营支出 (operational expenditure, OPEX),满足现代云计算环境下日益复杂的业务需求,进一步增强网络的可靠性与弹性^[3]。

NFV 概念最早由 ETSI 于 2012 年提出,并在此基础上成立了网络功能虚拟化行业规范组 (network function virtualization industry specification group, NFV ISG),以推动 NFV 相关技术的标准化与生态系统建设。自成立以来,NFV ISG 持续发布涵盖 NFV 技术机遇、应用场景、架构框架等多个方面的重要技术文档,为 NFV 的理论研究、技术发展及产业推广提供了重要的标准化指导和技术支撑。ETSI 定义的 NFV 架构框架如图 1 所示。

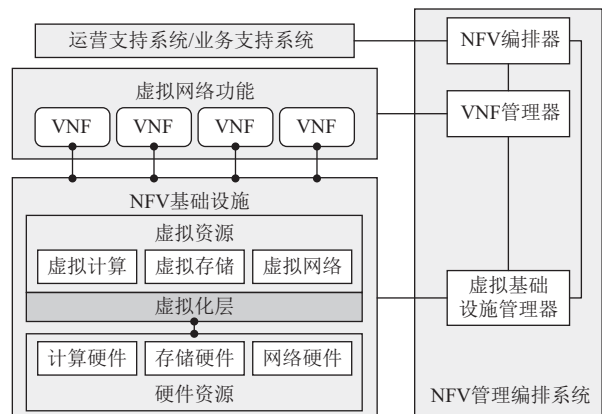


Fig. 1 Framework of NFV architecture

图 1 NFV 架构框架

NFV 架构主要由网络功能虚拟化基础设施 (network function virtualization infrastructure, NFVI)、NFV 和管理与编排 (management and orchestration, MANO)

三大核心模块构成^[20]。其中, NFVI 由物理硬件和虚拟化层组成, 为 VNF 的运行提供基础计算、存储及网络资源支持, 并通过虚拟化技术实现资源的高效管理和共享。VNF 作为网络功能的虚拟化实例, 涵盖了传统网络中的路由、防火墙、负载均衡等核心功能, 并具备根据流量需求动态调整的能力, 从而提高网络的灵活性和适应性。MANO 则负责整个 NFV 环境的资源协调与管理, 主要由 NFV 编排器(NFV orchestrator, NFVO)、虚拟网络功能管理器(virtualized network function manager, VNFM)和虚拟基础设施管理(virtualized infrastructure manager, VIM)三部分组成, 支持 VNF 的全生命周期管理, 确保资源的高效分配与网络功能的灵活调度。此外, 在电信运营场景中, 运营支撑系统(operation support system, OSS)和业务支撑系统(business support system, BSS)同样是关键组成部分。其中, OSS 主要负责服务提供、网络配置及维护, 而 BSS 则侧重于客户管理、账单处理和订单管理。

1.2 SFC 基本架构

SFC 的概念最初由互联网工程任务组中的(Internet Engineering Task Force, IETF)SFC 工作组提出, 旨在基于 SDN 的原理设计一种灵活的 SFC 架构。其核心理念是将逻辑服务功能链与数据平面的实际流量转发操作分离, 并规定如何根据预定义的 SFC 策略将流量在部署的服务功能(service function, SF)之间进行有序引导^[5]。通过控制平面与数据平面分离的设计原则, SFC 架构实现了对底层网络拓扑的独立性, 从而大大提高了网络管理的灵活性和可扩展性^[14]。此外, ETSI 工作组致力于将 IETF 定义的 SFC 组件整合至 NFV 架构中, 提出了 VNF 转发图(VNFFG)的概念, 用于描述服务功能链及其业务^[20]。VNFFG 与 SFC 定义一致, 描述了 VNF 的顺序及路径。图 2 展示了 NFV/SDN 架构下的 SFC 体系结构。

SFC 的部署涉及服务层、管理与编排层、控制层和转发层。

- 1) 服务层。负责逻辑上连接各 VNF, 并由 OSS/BSS 管理用户请求和策略配置。
- 2) 管理与编排层。以 NFVO 为核心, 负责虚拟资源的生命周期管理及编排, 提供统一接口管理 SFC 的实例化与配置。
- 3) 控制层。以 SDN 控制器为核心, 将相应的 SFC 策略解析为具体的服务功能路径(service function path, SFP), 并依据 VNF 状态及链路负载情况实时调整路径。

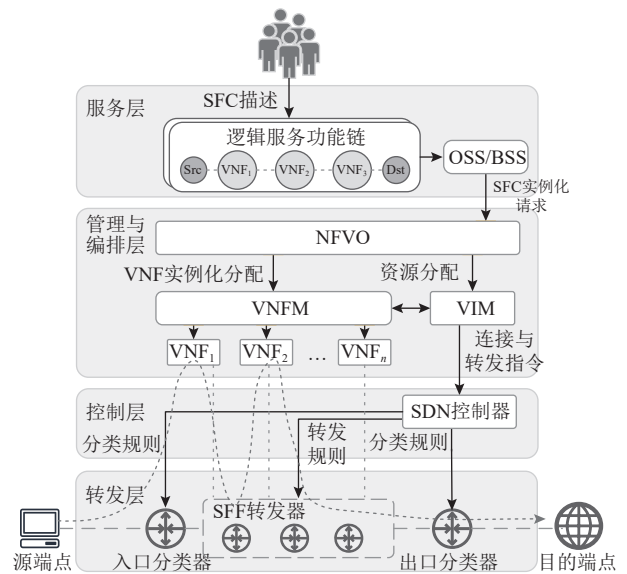


Fig. 2 SFC architecture in NFV/SDN architecture

图 2 NFV/SDN 架构下 SFC 体系结构

4) 转发层。由 SFC 分类器、SF 转发器、SF 及 SF 代理等组件组成, 负责数据包分类、路径标识及流量转发, 确保数据按既定 SFP 顺序通过各服务节点。

1.3 SFC 部署流程

SFC 的部署是一个复杂且多阶段的过程, 涵盖了从服务描述到功能放置与调度的全过程。每个阶段都涉及网络资源管理、QoS 保障和系统灵活性等方面的挑战。为了实现高效部署与优化, 可将 SFC 部署划分为以下 5 个主要阶段: 服务功能链的描述、构建、放置、调度及重配置, 如图 3 所示。

1) SFC 的描述

描述是 SFC 部署的起始阶段, 定义了 SFC 请求的功能属性和资源需求, 包括每个 VNF 的处理功能、

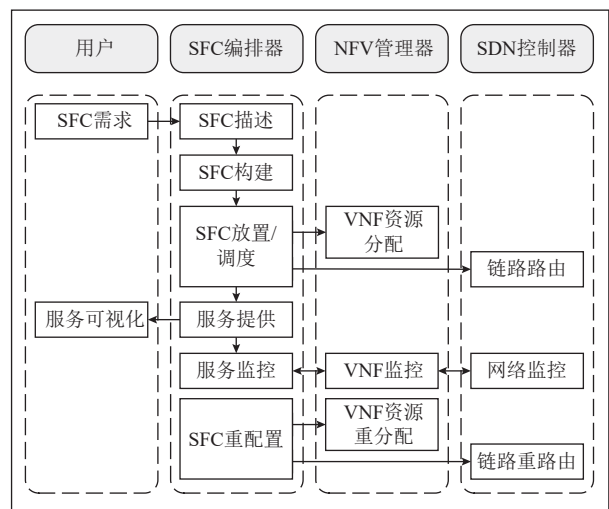


Fig. 3 Deployment workflow of SFC

图 3 SFC 部署流程

资源需求(CPU、内存、存储等)以及服务质量约束(如时延、带宽等)。

传统的描述方法如网络服务描述语言(Web services description language, WSDL)^[21]虽然在静态网络服务中表现良好,但在动态网络中表现出灵活性不足。为解决这一问题,IETF提出了基于(yet another next generation, YANG)数据模型^[22]的解决方案。YANG模型凭借其灵活性和层次化配置能力,逐渐成为主流方案^[22]。例如,文献[23]基于YANG模型提出了上下文树语法,通过定义弹性服务描述框架,显著提升了SFC在分布式环境中的可操作性。此外,云应用拓扑与编排规范(topology and orchestration specification for cloud applications, TOSCA)^[24]也系统地定义了服务功能的类型及其相互关系,在ETSI和开放网络自动化平台(open network automation platform, ONAP)中得到了广泛应用。

2) SFC的构建

SFC的构建涉及将多个VNF按照特定顺序组织成完整的服务链。尽管各VNF在功能上相对独立,但它们之间存在严格的处理顺序依赖关系,这使得VNF的排序成为服务链构建的核心问题。为了解决这些依赖关系,研究者们提出了多种有效的服务链组合算法。文献[23]设计了一种基于贪婪策略的算法,通过最小化数据流量来优化VNF的连接顺序。

此外,SFC的构建还涉及并行服务功能链的设计,多个服务功能可以并行处理数据流,从而提升系统吞吐量和响应速度。文献[25]通过分析VNF间的依赖关系,提出了将可并行的VNF部署在同一服务器上的方法以降低开销。考虑到服务器资源的有限性,且并行处理可能引发额外开销及数据包沉积等问题。相较于构建完全并行或串行服务链,文献[26]提出了PPC算法,该算法能够在确保性能的前提下,达成合理的部分并行化部署,为服务链构建中的并行优化提供了更具适应性与有效性的解决方案。

3) SFC的放置

作为SFC部署过程中的关键环节,SFC放置问题主要解决VNF的资源分配和网络部署。该问题可分为VNF节点放置和虚拟链路放置2个子问题^[17]。节点放置需要在物理基础设施中为每个VNF选择合适的服务器,而虚拟链路放置则负责在物理基础设施中构建连接这些VNF的网络路径。由于涉及多维资源约束和性能目标,VNF放置通常被证明是NP难问题^[27]。

从优化目标来看,现有研究集中在成本、时延、

可靠性及能效优化等方面^[17]。从算法来看,VNF放置问题的求解经历了从确定性算法到智能优化方法的发展历程。传统求解方法主要包括数学规划、启发式算法和元启发式算法3类。在数学规划方面,文献[28]将VNF放置建模为混合整数线性规划(mixed integer linear programming, MILP)问题,通过引入二进制变量表示VNF与物理节点的映射关系,并设计线性约束描述资源限制和QoS需求。文献[29]提出了一种区分等级的可生存服务功能链映射方法,将可生存SFC映射建模为MILP问题,并分别为关键SFC和普通SFC设计了主备路径构建算法与失效重映射算法。虽然MILP方法能获得最优解,但计算复杂度随网络规模呈指数增长,难以应用于大规模网络。为降低计算复杂度,研究者提出了多种启发式算法。文献[27]将VNF放置问题近似表述为一个整数线性规划(integer linear programming, ILP)模型,并采用一种启发式算法来解决该问题。面对动态网络环境带来的新挑战,研究者开始探索基于强化学习的智能优化方法。通过建立环境状态、动作空间和奖励机制,强化学习能够自适应学习部署策略,为VNF放置问题提供了新的解决思路。

4) SFC的调度

SFC调度阶段主要解决VNF的执行顺序和资源分配问题,其核心目标是在满足QoS要求的前提下,最大化系统吞吐量或最小化服务完成时间。文献[30]将VNF调度问题形式化为车间作业调度模型JSP(job shop problem),为后续研究奠定了理论基础。为降低服务完成时间从而使运营商能够服务更多客户,文献[31]将SFC调度与流量路由问题联合构建为MILP模型,并且提出了一种基于遗传算法的启发式算法以获取局部最优解。

5) SFC的重配置

SFC重配置通过动态调整已部署的服务链,应对网络环境的变化与业务需求的动态调整^[32-33]。不同于静态部署,重配置需在服务连续性保障的前提下优化资源利用率与服务性能^[34]。重配置的触发条件通常包括:①网络负载变化(如流量激增或资源瓶颈)^[32];②基础设施状态改变(如节点故障或拓扑更新);③服务需求调整(如QoS要求变化或业务优先级调整)^[33-34]。

从本质上看,SFC部署问题是一个多阶段、多目标、多约束的资源优化问题,其复杂性主要源于多个方面:物理网络中计算、存储、网络资源的异构性与VNF多样化需求构成了多维约束条件;网络状态和

服务需求的动态变化与不确定性使静态优化方法难以适应实际环境;资源利用率、能耗效率、服务质量和运营成本等多个优化目标之间存在冲突关系需要权衡;网络规模和服务链数量增长导致的决策空间组合爆炸使穷举搜索变得不切实际;此外,SFC部署各阶段之间的相互影响与密切关联形成了阶段耦合性,前一阶段的决策直接影响后续阶段的优化空间,因此理想的优化策略应当考虑全局最优而非各阶段的局部最优。

面对 SFC 部署各阶段的挑战,传统优化方法和智能优化方法各有优劣。整数线性规划和混合整数线性规划等精确求解方法能够提供全局最优解,保证解的质量,但其计算复杂度极高(通常为 NP 难),随着网络规模的增长呈指数级上升,导致其在大规模网络中求解时间过长,难以满足实时部署需求^[28]。此外,这类方法对问题模型的依赖性强,当网络环境发生变化时,需要重新构建模型并求解,缺乏灵活性。相较之下,启发式和元启发式算法(如贪婪算法、遗传算法、模拟退火等)通过特定的搜索策略,能够在可接受的时间内找到近似最优解,计算复杂度较低,适用于中等规模的网络环境^[27,31]。这类方法在求解效率和解的质量之间取得了平衡,但往往依赖于人

为设计的启发规则,在面对复杂多变的网络环境时,这些固定规则可能难以适应网络的动态变化,缺乏自主学习和优化能力。

近年来,强化学习等智能优化方法逐渐应用于 SFC 部署问题。强化学习方法之所以适用于 SFC 部署问题,源于其与问题本质的匹配性。SFC 部署本质上是一个序列决策问题,每个 VNF 的部署位置和路径选择都会影响后续决策空间,符合强化学习的马尔可夫决策过程(Markov decision process, MDP)框架^[35],强化学习通过智能体与环境交互,学习长期最优的决策策略,能够捕获决策间的依赖关系^[7]。同时,SFC 部署面临高维状态空间(包括网络拓扑、资源分布、流量需求等)和深度强化学习结合深度神经网络的表征能力,能够有效学习复杂的状态-动作映射,无需显式建模每个状态转移概率。此外,网络环境的动态变化可自然映射为强化学习框架中的环境变化,强化学习智能体通过持续交互和奖励反馈,能够自适应调整策略,适应环境动态性。SFC 部署通常需要平衡多个相互冲突的目标,强化学习通过奖励函数的设计可以隐式处理多目标优化问题,学习在各目标间取得平衡的策略。表 2 总结了不同优化方法在计算复杂度、求解质量、适应性等方面的对比。

Table 2 Comparison of SFC Optimization Methods

表 2 SFC 优化方法对比

方法	计算复杂度	求解质量	适应性	适用场景
ILP/MILP	高(指数级时间复杂度)	精确解	低(环境变化需重新建模求解)	小规模静态网络
启发式算法	中等(多项式时间复杂度)	近似最优解	中等(依赖预设规则,难以自适应调整)	中等规模静态网络
强化学习	训练阶段高,决策阶段低	随经验积累逐渐接近最优	高(能通过交互学习适应环境变化)	动态大规模网络

从表 2 可知,强化学习在应用于 SFC 部署时也面临挑战:训练过程需要大量环境交互,模拟器的构建和参数调优复杂,奖励稀疏问题可能导致学习效率低下。

1.4 基于强化学习的 SFC 部署框架

深度强化学习已在自然语言处理、机器人控制、决策博弈等领域展现出显著价值^[36],AlphaGo 的成功应用^[37]更是彰显了其卓越效能。强化学习的核心在于控制策略的学习与优化,通过智能体与环境的持续交互,实现累积回报最大化^[8]。马尔可夫决策过程作为强化学习的理论基础,为智能体的交互机制和策略优化提供了严格的数学支撑,使得复杂的 SFC 部署决策问题能够在该框架下得到有效表征和求解。

强化学习算法根据其核心机制可分为基于值函数的方法、基于策略的方法和基于值与策略的方法

等类别^[38-39]。基于值函数的方法如 Q-Learning、深度 Q 网络(deep Q-network, DQN)、双深度 Q 网络(double deep Q-network, Double DQN)等,通过估计状态-动作对的价值来选择最优动作,适合处理离散动作空间,但在连续动作空间中面临扩展性挑战。基于策略的方法如 REINFORCE、策略梯度(policy gradient, PG)、信任域策略优化(trust region policy optimization, TRPO)等直接优化策略函数,可以有效处理连续动作空间,具有较好的收敛性能。基于值和策略的方法如 AC(actor-critic)、深度确定性策略梯度(deep deterministic policy gradient, DDPG)、近端策略优化(proximal policy optimization, PPO)等,结合了基于值函数的方法和基于策略的方法的优势,通过 Actor 网络生成动作以及通过 Critic 网络评估动作价值,在复杂任务中表现出色。近年来,随着深度学习技术的发展,多

智能体方法如多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MADDPG)、多智能体近端策略优化(multi-agent proximal policy opti-

mization, MAPPO)进一步扩展了强化学习在复杂系统中的应用场景。表3详细对比了各类强化学习算法的特点及优劣。

Table 3 Comparison of Reinforcement Learning Algorithms

表3 强化学习算法对比

类别	算法	特点及优势	劣势
基于值的方法	Q-Learning	通过更新Q值表学习最优策略。算法实现简单, 适合离散动作空间。	对于高维状态/动作空间表现较差, 需要探索策略。
	DQN	使用神经网络逼近Q值函数, 适合高维状态空间, 结合经验回放缓解样本相关性问题。	训练时间较长, 难以处理连续动作空间, 容易出现过估计问题。
	Double DQN	改进DQN, 通过双网络缓解Q值过估计问题。算法更稳定, Q值估计更准确, 收敛效果好。	实现复杂度比DQN略高。
	TD (temporal difference) Learning	基于时序差分的学习方法, 结合蒙特卡洛与动态规划优点, 适用于在线学习。	对初始值敏感, 需要手动设计探索策略。
基于策略的方法	REINFORCE	基本的策略梯度算法, 实现简单, 可处理连续动作空间	高方差、样本效率低、收敛速度慢。
	PG	基于策略直接优化目标函数, 使用梯度更新策略; 简单易实现, 可直接处理连续动作空间。	高方差、收敛速度慢、可能陷入局部最优。
	TRPO	信任域策略优化, 保证策略更新单调改进, 稳定性高, 适合大规模复杂任务。	实现复杂、计算开销大、实现难度大。
基于值和策略的方法	AC	结合策略梯度和值函数, 减少方差, 提高收敛速度; 适合连续动作空间问题。	Actor 和 Critic 网络需要同时训练, 设计和调参复杂。
	A2C	同步版本的AC算法, 使用优势函数A代替Q值或V值。	性能不如异步版本, 样本利用效率低。
	A3C	异步版本的AC, 使用多线程并行更新全局网络。	实现复杂, 异步更新可能导致不稳定性。
	DDPG	基于Actor-Critic的离线算法, 高效处理连续动作空间任务, 样本效率高。	对超参数敏感, 探索能力不足, 可能不稳定。
	PPO	改进TRPO, 使用简单但有效的剪辑机制来限制策略更新幅度; 实现简单、稳定性好、收敛速度快; 适合大规模任务。	超参数较多, 可能需要较多调试。
多智能体方法	MADDPG	DDPG的多智能体扩展, 针对每个智能体使用独立Actor和共享Critic。	对环境建模要求高, 协作机制复杂, 可能不稳定。
	MAPPO	PPO算法的多智能体扩展, 适用于多智能体协作或竞争环境; 可扩展性好、训练稳定。	计算开销较大, 通信和协作机制设计复杂。

图4展示了针对SFC部署问题的MDP建模方法。为应对请求随机到达与离开的网络动态变化, 引入时隙的概念, 在每个时隙内按既定流程顺序执行SFC操作。具体而言, 状态通过多维向量形式全面表示网络在各时隙下的运行状况, 确保了对实际环境的精确描述; 动作定义则明确了VNF的部署策略;

奖励函数的设计同时考虑了运营商资源效率和客户QoS需求, 实现了多目标优化; 而状态转移则刻画了网络在不同时隙间的动态演变过程, 为制定更加高效的部署策略提供了依据。这一模型框架有效地为SFC部署提供了形式化表述。

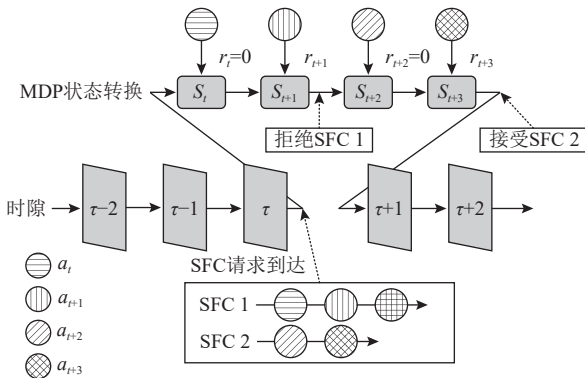


Fig. 4 MDP modeling of SFC deployment process

图4 SFC部署过程的MDP建模

基于上述MDP建模, 可进一步构建基于强化学习的SFC部署框架, 如图5所示。该框架的核心在于智能体与环境的交互机制, 通过不断循环迭代实现部署策略的动态优化。环境包括网络拓扑、实时流量状态及资源可用性等动态要素, 可作为智能体感知和决策的依据。智能体在每个时隙内根据当前环境状态选择部署动作, 即为VNF制定具体部署方案, 这一决策直接影响系统性能。环境则通过奖励信号量化部署方案的效果, 反馈给智能体用于优化其策略。智能体依据累计奖励不断调整行为, 从而实现网络资源利用率与QoS之间的优化平衡。

通过上述交互过程, 基于强化学习的SFC部署框架能够适应复杂网络环境的动态变化, 有效提升

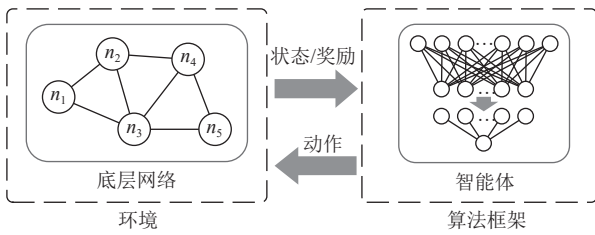


Fig. 5 SFC deployment framework based on reinforcement learning

图5 基于强化学习的SFC部署框架

SFC的部署效率与性能。该框架为应对多目标优化和资源分配问题提供了新思路^[7]。本文后续章节将重点探讨强化学习技术在SFC放置、调度与重配置中的具体应用,系统分析相关研究进展,并展望未来发展趋势。

2 基于强化学习的SFC放置研究

基于强化学习的SFC放置问题研究经历了从简单到复杂、从单一到协同的发展历程。早期研究主要采用传统强化学习方法解决小规模场景下的放置优化问题。随着深度学习技术的发展,研究者逐渐将深度强化学习引入到SFC放置决策中,以应对大规模复杂场景。近年来,图神经网络和多智能体技术的引入进一步提升了放置决策的效率和准确性。本节将从技术演进的角度详细阐述基于强化学习的SFC放置研究进展。

2.1 基于传统强化学习的方法

传统强化学习通过智能体与环境交互来学习最优策略,其简单有效的特点使其成为早期解决SFC放置问题的重要方法。

文献[40]提出了一种基于Q-Learning的动态SFC放置方案。该方法通过构建状态空间,描述网络服务功能的执行时间、节点资源使用率、服务链时延和链路带宽等信息;同时设计奖励函数以平衡资源利用与负载均衡目标。智能体通过不断地试错学习,选择资源利用率低的节点来构建服务路径,从而实现节点负载均衡和资源高效使用。然而,该方案在大规模网络场景下由于状态空间的急剧膨胀,容易出现维数灾难问题,导致学习收敛速度下降。

为缓解上述问题,文献[41]提出了一种结合强化学习和启发式算法的混合解决方案——QLFHM算法。该算法将强化学习模块用于输出满足特定要求的最短路径,随后通过负载均衡模块在这些路径中挑选最优解,进一步提高了放置策略的效率与性

能。实验结果表明,该算法在动态环境下的放置成功率高于传统启发式方法,并能有效降低计算时间。

尽管这些研究充分展示了传统强化学习在SFC放置问题中的潜力,但其局限性也较为显著。Q-Learning等值函数型算法依赖于离散化的状态和动作空间,在处理大规模连续状态空间时效率较低^[42]。随着网络规模和复杂度的增长,传统强化学习方法的局限性愈加凸显。

2.2 基于深度强化学习的方法

随着深度学习技术的发展,强化学习与深度学习的结合,即深度强化学习,在处理复杂决策问题方面展现出了显著优势。深度强化学习不仅能够有效解决传统强化学习在高维状态空间中面临的维数灾难问题,还能通过神经网络的强大特征提取能力,直接从原始数据中学习出有效的放置策略。近年来,基于深度强化学习的SFC放置方法取得了广泛关注。

1) 基于DQN的方法

深度Q网络是深度强化学习的经典方法之一,其通过深度神经网络逼近Q值函数,实现了从状态到动作价值的端到端学习。文献[7]将深度强化学习引入SFC放置领域,提出了NFVdeep方法。该方法将SFC放置问题建模为MDP,在状态空间设计上综合考虑网络资源利用率、当前SFC部署状态及VNF特征,构建了完整的网络状态向量。在动作空间定义方面,采用整数编码表示VNF放置策略。奖励函数设计则通过加权计算接受请求收益与资源开销的差值,有效平衡了服务提供商和用户的利益。基于此MDP框架,NFVdeep采用序列化-回溯方法处理离散动作空间,并利用策略梯度法优化训练效率与收敛性。实验表明,该方法在请求吞吐量和运营成本方面均优于传统基准算法。

文献[43]从QoS保障角度出发,将SFC放置建模为参数化动作马尔可夫决策过程,提出了基于策略梯度的DRL-QOR方法。该方法以最大化用户体验质量为目标,同时满足QoS约束,采用深度神经网络近似Q函数。不同于逐步放置方式(每一步仅放置1个VNF),DRL-QOR采用单步放置策略,在一个决策步骤内生成所有VNF的放置方案。

文献[44]关注NFV环境下在线服务供应面临的资源分配挑战,提出了一种基于DQN的自适应资源分配算法。与NFVdeep不同,该方法考虑了服务生命周期的概率特性,通过随机过程建模捕捉服务请求的动态性。文献[45]针对具有优先级约束的SFC放置问题,提出了基于DQN的自适应映射方法

ADAP。该方法首先将多个具有优先级约束的 SFC 放置问题形式化为多步骤整数线性规划问题,然后设计了 2 种基于贪婪策略的低复杂度算法,并利用 DQN 结合 ϵ -贪婪策略实现自适应调度。

文献 [46] 针对边缘云环境中的服务功能链部署挑战,提出了 SDRM-EC-DPR 算法,旨在最大化在线 SFC 部署的收益。与前述工作不同,该研究在边缘云融合环境下全面考虑了排队延迟、计算延迟和通信延迟的综合影响,并基于市场供求规律构建了统一的部署成本模型,消除了边缘云环境中设备的异构性。在算法设计上,SDRM-EC-DPR 基于对决型双深度 Q 网络 (dueling double deep Q-network, D3QN) 框架,进一步融合了优先经验回放机制和随机网络蒸馏技术,提升了学习效率并避免了局部最优陷阱。

2) 改进的 DRL 算法

随着研究的深入,学者们开始关注深度强化学习框架的创新和改进。针对大规模动作空间引发的维度灾难问题,文献 [47] 提出了具有增强探索能力的 DDPG 算法。该算法通过引入启发式算法来求解可行解,从而有效确保动作的可行性并提升算法效率。文献 [48] 则提出一种基于改进的 DDGP 算法,基于价值差异的探索方法对贪婪策略进行扩展,以此来平衡智能体在动作选取时探索与利用之间的关系。同时,该算法设计了双重经验回放池机制,从而提高训练效率和收敛速度。此外,文献 [49-50] 在 VNF 放置和路由问题研究中引入在图像处理 and 机器翻译中广泛应用的注意力机制。通过将注意力机制应用于 DDPG 框架中的智能体动作建模,使模型能够自动关注对智能体决策具有重要影响的邻居节点,有效降低了对不相关节点的计算开销,显著提升了模型训练效率。实验结果表明,与传统 DQN 方法相比,基于注意力机制的方法在网络效用、时延控制、成本优化和请求接受率等多个性能指标上均取得了显著提升。文献 [51] 提出基于增强 E-DDPG 算法通过重构缓冲重放机制(包括多样性缓冲重放和高误差缓冲重放),显著提升了智能体的训练效率。文献 [52] 针对边缘智能赋能的工业物联网中动态 SFC 编排问题,提出了基于 SAC (soft actor-critic) 的深度强化学习方法。该方法充分考虑了工业物联网服务请求的时延约束、时变特性和异构性,通过优化 SAC 算法提升了模型的稳定性和收敛性。

3) 可靠性导向的 SFC 放置方法

可靠性和容错性作为优化目标的 SFC 放置问题逐渐受到研究者的关注,主要聚焦于 SFC 放置过程

中的故障容错、服务可用性保障和系统稳定性维持。文献 [53] 针对在线容错 SFC 放置问题,提出了基于双深度 Q 网络的 DDQP 算法和 DDQP+算法,分别负责决策运行时和备用服务实例的放置位置。为了提供全面的容错放置解决方案,研究者同时提出了 5 种渐进的资源预留方案,从零预留到全量预留构建了完整的容错策略谱系。文献 [54] 提出了一种基于 DRL 的优先级感知算法,通过计算 SFC 和网络节点的优先级来确定备份方案,利用深度强化学习算法选择 VNF 的放置和备份位置,提升了系统的可靠性和服务质量保障能力。文献 [55-56] 针对 SFC 可用性、运营成本 and 能耗等因素的 SFC 放置问题,提出一种基于强化学习的方法,通过优化节点选择和冗余策略,有效平衡了服务可用性与资源利用效率。

4) 多目标优化 SFC 放置方法

SFC 放置通常需要同时考虑多个优化目标,如时延、资源利用率、能耗等。基于深度强化学习的多目标优化方法试图在这些冲突的目标之间寻找帕累托最优解。研究者们提出了多种策略来处理多目标问题,包括基于权重的方法^[57]、基于帕累托前沿^[58]的方法等。文献 [59] 提出了一种切比雪夫辅助的 Actor-Critic SFC 放置算法,创新性地将非线性切比雪夫函数与 Actor-Critic 算法相结合以探索帕累托前沿。该方法首先将单目标混合整数线性规划模型扩展为多目标模型;其次设计了基于指针网络 (pointer network, PN) 的 DRL 模型来处理大规模部署问题。该算法无需人为干预即可逼近帕累托前沿,通过选择不同权重来改善边缘计算资源使用和服务接受率。仿真结果表明,该算法在生成非支配前沿方面表现卓越,可有效规避边缘资源拥塞,在不同负载场景下实现高服务接受率。

5) 并行 SFC 放置方法

针对串行 SFC 时延性能和处理效率方面的局限性,研究者们提出了基于强化学习的并行 SFC 放置方法。文献 [60] 针对移动边缘计算中 SFC 放置问题,提出了一种基于深度强化学习的并行化 VNF 动态 SFC 放置算法 DSPPV。该算法使用 A3C (asynchronous advantage actor-critic) 算法作为 DRL 框架来解决分布式网络中的并行 SFC 放置问题。DSPPV 还包含一种提取初始化 VNF 分布的策略,通过配置 VNF 队列网络,预测 A3C 中智能体最合适的动作,从而提高未来请求的接受率。文献 [61] 为解决数据中心网络中异构带宽需求下的 SFC 放置问题,提出了基于 DRL 的 SPC3D 算法。该算法将原始 SFC 划分为多个子

SFC, 通过分析 VNF 间的依赖关系实现 SFC 的并行化处理, 有效降低了服务时延。实验结果表明, 与传统方法相比, SPC3D 在处理大规模异构带宽需求时表现出更好的性能和稳定性。

基于深度强化学习的 SFC 放置方法通过深度神经网络的强大表征能力, 成功解决了传统方法在处理高维状态空间和复杂决策时的局限性。从最初的 DQN 到改进的 DDPG, SAC 等算法, 以及在多目标优化、可靠性保障、时延优化、资源优化和并行放置等方面的深入研究, 深度强化学习展现出显著的优势。特别是在处理动态网络环境、异构资源需求和多维度优化目标等复杂场景时, 基于深度强化学习的方法能够自适应地学习最优放置策略, 在性能、效率和适应性等方面均优于传统方法。然而, 这类方法仍存在训练开销大、可解释性不足等问题, 这也促使研究者探索将其与其他先进技术相结合的新方向。

2.3 图神经网络辅助的方法

SFC 部署问题通常涉及复杂的网络拓扑结构和多维资源约束, 如何高效地提取网络拓扑信息以指导部署决策是关键难题之一。图神经网络(GNN)因其在处理图结构数据中的独特优势, 越来越多地被集成到强化学习框架中, 用于优化复杂网络环境下的 SFC 部署问题。GNN 通过节点嵌入和边特征聚合的方式, 能够充分捕获网络节点和链路间的关联性, 生成具有全局上下文信息的网络表示。

文献 [9] 针对传统基于深度神经网络的 SFC 放置方法在利用网络拓扑信息方面存在的不足, 提出了一种基于编码器-解码器架构的 GNN 模型。该模型的编码器通过结合门控循环单元(gated recurrent unit, GRU)的 GNN 对网络拓扑进行特征提取, 解码器则基于编码得到的网络表征逐步选择放置节点并评估 VNF 的处理概率, 从而生成完整的 SFC 部署路径。文献 [62] 在解决 VNF 放置问题时, 创新性地将深度强化学习与 GNN 相结合, 提出了 DeepOpt 方案。该方案在 GNN 结构设计上采用了图网络(graph network, GN)^[63] 来实现节点和边缘属性的动态更新与聚合, 有效提升了对网络拓扑结构的感知能力。进一步地, 文献 [64-66] 提出的算法不仅将底层物理网络建模为图结构, 还将服务请求表示为有向无环图, 通过 GNN 分别对 SFC 请求和底层网络进行特征提取, 实现了请求和底层资源的双重 GNN 表征。

综上所述, 基于 GNN 的方法在 SFC 部署问题中展现了强大的能力, 不仅能够有效捕获复杂网络拓扑的全局信息, 还能够通过与强化学习的深度融合,

进一步优化部署策略。

2.4 多智能体协同放置方法

多智能体强化学习方法因其能够通过多个智能体的协同决策解决大规模复杂问题, 成为 SFC 部署研究的一个重要方向。在多智能体场景中, 每个智能体可以独立感知环境、执行动作, 并通过相互协作完成复杂任务。与单智能体方法相比, 多智能体方法不仅能够更高效地处理分布式网络环境中的部署任务, 还能更好地应对大规模网络中资源分配的动态性和异构性。本节从智能体设计和协同机制角度探讨多智能体强化学习在 SFC 放置中的典型应用与研究进展。

1) 基于 SFC 请求分配多智能体的方法

在网络中, SFC 的放置往往涉及多个请求, 而不同请求的 QoS 需求和资源约束各异。基于请求分配的多智能体方法通常采用“1 个智能体对应 1 个 SFC”的设计模式, 每个智能体负责处理单个 SFC 的放置和优化任务。这种设计可以有效地实现请求级别的分布式决策, 提高放置效率和系统灵活性。

文献 [10] 提出了一种针对物联网中资源分配问题的多智能体强化学习方法。该方法中, 每个智能体管理一个 SFC 请求, 并通过 2 阶段决策优化 SFC 的放置和资源分配。在第 1 阶段, 智能体集中选择需要服务的物联网设备; 在第 2 阶段, 通过多智能体协作优化虚拟服务请求的映射和资源分配。文献 [67] 利用马尔可夫博弈建模多用户网络资源竞争问题, 并通过 MADDPG 算法求解竞争的纳什均衡, 从而优化了 SFC 的资源分配和放置策略。该方法在保证 QoS 的同时, 均衡了多用户之间的资源竞争。此外, 文献 [68] 提出了一个 2 阶段基于 GCN 的深度强化学习框架, 用于解决多数据中心网络中的 SFC 嵌入问题。第 1 阶段将负载转移过程建模为 MDP, 通过基于 GCN 的 PPO 算法实现 SFC 请求的负载均衡; 第 2 阶段利用多智能体框架, 采用基于 GCN 的多智能体 MAPPO 算法进行精细化放置优化。该框架通过 2 阶段协同优化提升了放置效率和资源利用率。

2) 基于节点协同的分布式放置方法

考虑到网络基础设施的分布式特性, 基于节点协同的多智能体方法在每个网络节点分配独立的智能体, 通过智能体间的信息交互和协同决策完成 SFC 放置。这种方法充分利用了分布式系统的优势, 智能体通过消息传递机制共享局部观察信息和决策意图, 实现资源分配和路径规划的全局优化。

文献 [69] 针对成本和时延敏感的 SFC 放置问题,

提出了多智能体深度强化学习算法 MADRL-P&R。该算法将 SFC 放置问题分解为 2 个子任务,通过设计联合奖励和内部奖励机制实现任务间的协同优化。同时,引入参数迁移的模型重训练方法,有效提升了算法对网络拓扑变化的适应能力。文献 [70] 聚焦于多域算力网络中的 SFC 映射问题,提出了基于数字孪生辅助的 MADDPG 框架。该框架将 SFC 映射建模为长期优化问题,综合考虑计算能力、链路带宽、时延性能和 VNF 排序等多维约束,以降低能耗、端到端时延和 VNF 重新放置成本为优化目标。通过引入数字孪生技术,该方法提升了决策的准确性和实时性。

3) 面向多目标优化的协同决策方法

面向多目标优化的多智能体方法采用“1 个智能体对应 1 个优化目标”的范式,通过多个智能体的协同决策来实现多目标之间的动态平衡。

文献 [71] 针对现有研究多集中于单目标或静态场景优化,且在动态多目标优化方面存在不足的问题,提出一种基于 GCN 的多智能体深度强化学习方法。该方法构建了包含物理网络拓扑特征提取的 GCN 神经网络和 Actor-Critic 模型,通过多智能体强化学习框架,使不同智能体对应不同奖励函数,实现对动态环境变化的感知与 SFC 放置策略的调整。实验验证表明,该方法在动态网络环境下能够提升放置性能,并实现多目标之间的有效平衡。

多智能体协同放置方法的研究表明,通过合理的智能体划分和协同机制设计,可以有效提升 SFC 放置的决策效率和优化效果。然而,当前研究仍存在智能体通信开销大、训练收敛性不稳定等问题。

为了系统地总结和对比基于强化学习的 SFC 放置方法,表 4 从强化学习算法、优化目标、VNF 放置策略以及虚拟链路放置策略等维度进行了详细归纳。

从优化目标来看,早期研究主要关注单一目标优化,如运行成本或时延;随着研究深入,越来越多的工作开始考虑多目标优化问题,如资源利用率、负载均衡、能耗等。在放置策略方面,不同方法根据具体场景和优化目标采用了节点序号、路由路径等多样化的策略选择方式。

3 基于强化学习的 SFC 调度研究

SFC 调度问题主要关注在已知网络资源和 SFC 请求的情况下,如何为每个 SFC 制定合理的资源分配方案和执行顺序,以达到系统性能的最优。与

SFC 放置问题相比,调度问题更加强调服务执行的时序性和资源的动态分配。特别是在现代网络架构中,随着 5G/6G、边缘计算等技术的快速发展,对服务响应时间的要求愈发严格,这使得时延优化成为 SFC 调度研究的核心目标之一。本节从优先级机制角度对基于强化学习的 SFC 调度研究进行分类讨论,主要分为基于任务特性的优先级和基于服务需求的优先级两大类。

3.1 基于任务特性的优先级

基于任务特性的优先级机制主要关注 SFC 或 VNF 本身的属性和特征,以任务的固有特性作为确定优先级的依据。这类方法主要聚焦于如何通过合理的调度顺序优化系统整体性能指标,如总完成时间、资源利用率或系统吞吐量等。

在移动边缘计算场景中,文献 [72] 将 SFC 调度问题建模为灵活作业车间调度问题,并提出了基于 DQN 的调度算法,以最小化总体调度时延为优化目标。该研究首先根据 SFC 中未处理的 VNF 数量为 SFC 分配优先级,具有更多 VNF 请求的 SFC 获得更高的处理优先级。其次,在动作匹配过程中也考虑了优先级因素,当多个动作满足条件时,会选择优先级更高的动作执行。这种多层次的优先级设计有效平衡了资源利用率与调度时延。

针对多资源约束下的 SFC 调度问题,文献 [73] 首先证明了该问题的 NP 难特性,随后设计了基于 DRL 的求解框架。在该框架中,智能体通过感知服务器资源使用状况和 SFC 状态来做出调度决策,选择最适合的服务器与 SFC 进行匹配。其奖励函数设计与未完成 SFC 数量建立关联,引导智能体朝着最小化平均流完成时间的方向优化。该研究在资源调度时允许暂停当前 VNF 的执行,优先处理可能带来更高收益的 VNF,实现了动态优先级调度。其次,将最短作业优先和最小资源优先等优先级启发式集成到 DRL 训练过程中,显著提高了训练效率。

文献 [74] 提出基于 DRL 的 VNF 协同调度框架,用于解决传统 SFC 批处理放置和调度方法在处理多样化和优先级任务时的不足。该框架创新性地提出了基于最短剩余时间优先规则的优先级处理机制,通过考虑 VNF 的排队等待时延和服务完成时延计算其优先级,并以最小化带优先级权重的平均时延为优化目标。同时,该文设计了基于优先级的多维资源分配策略,根据 VNF 在队列中的优先级位置分配计算资源比例,确保高优先级 VNF 获得更多资源和更少的排队时延。

Table 4 Comparison of SFC Placement Algorithms Based on Reinforcement Learning
表 4 基于强化学习的 SFC 放置算法对比

分类	放置算法	强化学习算法	优化目标	VNF 放置策略	虚拟链路放置策略
基础 DRL 算法	NFVdeep ^[7]	PG	运行成本、吞吐量	节点序号	
	DRL-QOR ^[43]	DQN	QoE	节点序号	
	DQN-AR ^[44]	DQN	运行成本、带宽资源	节点序号	链路序号
	ADAP ^[45]	DQN	带宽消耗、CPU 消耗	2 种基于贪婪的候选算法	重赋链路权重
	SDRM-EC-DPR ^[46]	Dueling Double DQN	收入、成本	边缘/云层设备选择	
	DDQP ^[53]	Double DQN	接受率、服务可用性	放置与备份节点序号	
	PA-DRL ^[54]	DQN	传输时延、网络负载平衡	放置与备份节点序号	
Actor-Critic 类算法	E ² D ² PG ^[47]	DDPG	接受率	二进制节点放置决策	重赋链路权重
	改进 DDPG ^[48]	DDPG	时延、部署成本	资源分配与放置节点	
	A-DDPG ^[49]	基于注意力的 DDPG	收入、成本	二进制节点放置决策	二进制链路放置决策
	A-DDPG ^[50]	基于注意力的 DDPG	收入、成本	二进制节点放置决策	二进制链路放置决策
	E-DDPG ^[51]	DDPG	时延	二进制节点放置决策	二进制链路放置决策
	Opt-SAC ^[52]	SAC	时延	节点放置概率	
	A2C ^[55]	A2C/PPO	可用性、能耗	节点序号与冗余实例数量	
	Cand-RL ^[56]	PPO	部署成本、能耗	候选节点序号与备份决策	最短路径
	Chebyshev-MODRL ^[59]	Actor-Critic	时延, 拥塞率	节点序号	最短路径
	DSPPV ^[60]	A3C	时延、资源消耗、吞吐量	节点序号	
SPC3D ^[61]	A3C	时延, 资源消耗、吞吐量	候选节点序号		
图神经网络 辅助方法	DeepOpt ^[62]	REINFORCE	部署成本	二进制节点放置决策	
	Kolin ^[64]	REINFORCE	部署成本、流量转发成本	节点放置概率	重赋链路权重
	GDRL-VNFP ^[65]	A3C	部署成本、流量转发成本	节点序号	最短路径
	DRL-D ^[66]	TD	长期平均收入	二进制候选节点放置决策	
多智能体协 同放置方法	MA-DQL ^[10]	DQN (独立学习)	接受率、部署成本	IoT 设备选择	
	MADDPG ^[67]	MADDPG (集中训练分散执行)	负载均衡、时延	节点序号	链路序号
	two-stage GCN-based DRL ^[68]	MAPPO (集中训练分散执行)	接受率、成本	路由路径	路由路径
	MADRL-P&R ^[69]	DDPG (独立学习)	成本、时延	节点智能体放置决策	链路智能体放置决策
	DT-assisted MADDPG ^[70]	MADDPG (集中训练分散执行)	能耗、时延、部署成本	二进制服务域决策	路由路径
GCN-MADDPG ^[71]	MADDPG (集中训练分散执行)	资源使用率、时延	节点序号		

针对空天地一体化网络中的动态 SFC 调度问题, 文献 [75] 提出了一种基于 DRL 的协同优化方法。该方法首先引入可重构时间扩展图来刻画空天地一体化网络的多层资源分布, 并设计了一种基于数据量的 SFC 优先级排序机制。同时, 其提出的 VNF 状态转换算法能够在资源冲突时智能调整处理顺序, 优先处理关键 VNF 以确保多个 SFC 同时满足时延约束。

3.2 基于服务需求的优先级

基于服务需求的优先级机制主要关注服务请求的差异化需求、服务质量要求和业务重要性, 以服务或用户的外部需求为确定优先级的依据。这类方法通常考虑不同服务之间的差异, 如时延敏感度、吞吐量需求、可靠性要求等, 以及不同用户群体的服务等

级协议(service level agreement, SLA), 旨在满足多样化的服务质量需求并实现资源的公平分配。

对于具有端到端时延要求的 VNF 调度问题, 文献 [76] 将其转化为 MDP 并提出强化学习调度算法。该算法在系统状态设计上全面考虑了 NFV 节点资源状态和 VNF 执行状态通过奖励函数中权重系数的设计实现了服务优先级的差异化处理。具体而言, 为高优先级(时延敏感)服务分配较大的权重系数, 为低优先级(时延不敏感)服务分配较小甚至为零的权重系数, 从而实现对不同优先级服务的差异化调度。

针对工业物联网场景中基于信息年龄(age of information, AoI)的 VNF 调度问题, 文献 [77] 提出了一种复合 A2C 深度强化学习框架。该框架同时考虑

了 VNF 的放置位置选择和执行顺序调度, 通过设计多维动作空间, 实现了放置成本、调度成本和平均 AoI 的联合优化。该研究引入了基于 AoI 的优先级调度机制, 将具有较高 AoI 值的服务请求赋予更高的处理优先级, 确保信息的实时性和新鲜度。同时, 将到达的服务请求分为当前到达服务、当前终止服务和活动服务 3 类, 针对不同类别采取差异化的优先级处理策略。

针对边缘计算中 VNF 放置和调度的公平性问题, 文献 [78] 提出基于 PPO 的协同优化算法。考虑到资源限制和任务多样性, 该算法旨在确保不同服务链之间的资源分配公平, 同时满足各自的端到端时延要求。该研究通过学习优先级策略函数, 为候选的映射和调度决策分配优先级值, 使得资源竞争激烈时能够公平地处理不同服务链的需求。该方法将多层感知器作为优先级分配规则的函数实现, 动态地根据服务链的状态确定 VNF 的映射和调度优先级。

针对医疗急救场景中的服务功能链调度问题, 文献 [79] 提出了一种基于业务紧急程度和服务区域的优先级分配机制。该研究将医疗业务分为远程紧急、本地紧急、远程普通和本地普通 4 类, 分别对其赋予不同权重值, 并以最小化总加权完成时间为优化目标建立服务功能链调度模型。与最小化最大完成时间目标不同, 该模型更关注高优先级服务的及时完成。该研究针对不同问题规模, 设计了匹配博弈算法和 Q-Learning 强化学习算法, 通过仿真实验验证了方法的有效性。

文献 [80] 针对车联网 (Internet of vehicles, IoV) 中边缘云服务功能调度问题, 提出了一种基于多智能

体强化学习的分层动态调度算法。该算法创新性地考虑了车辆节点移动导致的无线信号强度动态变化, 并在问题建模中明确设定了服务请求优先级约束。算法通过图注意力网络提取特征, 实现了服务请求选择和执行车辆选择的解耦, 在保证公平调度的同时, 为移动场景下的 SFC 调度提供了新的解决思路。

针对网络切片中 VNF 调度的时延问题, 文献 [81] 提出基于 DRL 的低时延 VNF 调度算法。该研究设计了一个基于复合规则的 D3QN 调度模型, 引入了多种 SFC 优先级选择策略, 包括基于截止时间、松弛时间、剩余处理时间与操作数比值、估计平均处理时间等多种优先级规则。在每个调度决策点, D3QN 算法能够根据当前网络状态自适应地选择最优的优先级规则, 而不是固定使用单一规则。实验结果表明, D3QN 调度算法在不同网络环境下均能保持优越性能。

尽管目前基于强化学习的 VNF 调度研究相对新颖、文献数量有限, 但已展现出良好的发展潜力和应用前景。表 5 总结了近年来基于强化学习的 VNF 调度算法特征。

从表 5 中可以看出, 当前研究主要集中在 5 个方面: 算法选择上, 从早期的 Q-Learning 发展到 DQN 及其变体 (如 D3QN, Double DQN), 再到策略梯度类算法 (如 PPO, A2C), 显示出强化学习算法在该领域的快速演进; 优化目标方面, 时延 (包括传输、传播、处理、排队时延) 和完成时间是最主要的关注点, 部分研究也考虑了资源利用率、接受率等多目标优化; 调度动作设计上, 主要表现为 2 种模式: 一是直接选择 VNF 的部署节点序号, 二是选择特定的调度算法或

Table 5 Comparison of SFC Scheduling Algorithms Based on Reinforcement Learning

表 5 基于强化学习的 SFC 调度算法对比

调度算法	强化学习算法	优化目标	VNF 调度动作	优先级机制
DQS ^[72]	DQN	传输、传播、处理、排队时延	MDC 序号	基于未处理 VNF 数量: 未处理 VNF 越多优先级越高
文献 [73]	DQN	完成时间	服务与节点序号	动态优先级: 可暂停当前 VNF 执行高收益任务; 集成最短作业优先和最小资源优先规则
VNF-CSWD ^[74]	Double DQN	传输、传播、处理、排队时延	VNF 节点序号	最短剩余时间优先: 根据排队时延和完成时延计算优先级
DRL-MSSNL-SAGIN ^[75]	Double DQN	接受率 (有限时延)	VNF 节点序号	基于数据量的 SFC 优先级: 资源冲突时按数据量升序排序处理
文献 [76]	Q-Learning	完成时间	VNF 节点序号	时延敏感度权重: 高时延敏感服务获得更大权重系数
CA2C ^[77]	Compound A2C	放置与调度成本、信息年龄	节点放置决策、功率和子载波分配	基于信息年龄: 信息年龄值越高优先级越高
PPO-MCTS ^[78]	PPO	处理时延	候选节点与调度策略	学习型公平优先级: 通过神经网络动态分配优先级确保公平性
文献 [79]	Q-Learning	总加权完成时间	服务与节点匹配	基于业务类型的静态权重: 按紧急程度和服务区域分为 4 级
文献 [80]	PPO	完成时间、资源利用率、接受率	服务与节点对	所有服务请求默认具有相同处理优先级, 通过强化学习算法动态调整处理顺序
D3QN ^[81]	Dueling Double DQN	接受率 (有限时延)	5 种调度算法的选择	复合规则自适应选择: 根据网络状态在 5 种优先级规则中选择最优方案

参数。应用场景则从传统数据中心扩展到边缘计算、车联网、工业物联网和空天地一体化网络,涵盖了多种新型网络场景,证明了强化学习在处理复杂约束条件下 VNF 调度问题的广泛适用性。

4 基于强化学习的 SFC 重配置研究

网络环境的动态性和复杂性对 SFC 的部署提出了更高要求。为适应网络流量波动、用户移动性变化和节点故障等情况, SFC 重配置成为保障服务质量的关键技术。基于强化学习的 SFC 重配置方法凭借其自适应学习和动态决策能力,为解决这一问题提供了新思路。本节将从 VNF 迁移、动态预测重配置、可靠性导向重配置等维度系统分析基于强化学习的 SFC 重配置研究进展。

4.1 使用 VNF 迁移的重配置方法

VNF 迁移是实现 SFC 重配置的核心机制之一,其主要目标是通过将 VNF 实例从资源受限或性能下降的节点迁移至更合适的位置,以维持服务质量并优化资源利用。然而, VNF 迁移过程涉及多个挑战: 1) 需要准确评估网络状态和服务质量变化趋势确定迁移时机; 2) 目标节点的选择需要综合考虑计算资源、网络带宽和服务时延等多维约束; 3) 迁移过程的执行需要最小化服务中断时间和资源开销。针对这些挑战, 研究者们提出了多种基于强化学习的创新解决方案。

文献 [82] 面向 5G 网络切片架构, 提出了一种基于深度强化学习的 VNF 智能迁移框架。该框架将 VNF 迁移问题建模为受限 MDP, 以最小化通用服务器平均能耗为优化目标, 同时将各切片的平均时延、缓存及带宽资源消耗作为约束条件。通过设计基于卷积神经网络(convolutional neural network, CNN) 近似行为值函数, 能够自适应学习最优的 VNF 迁移策略和 CPU 资源分配方案。实验结果表明, 该方法在满足切片服务质量需求的同时, 降低了基础设施能耗, 实现了资源利用率的有效提升。

针对数据中心网络中 SFC 的动态调整问题, 文献 [83] 考虑到网络请求波动导致初始部署无法满足用户资源需求的情况, 提出了一种智能服务功能链动态调整算法。该算法首先构建了以最小化长期调整成本和资源分布不平衡性为目标的整数线性规划模型, 并将其转化为 MDP。通过采用 Double DQN 架构, 算法能够根据网络状态和资源利用情况, 自适应决策 VNF 实例是否需要弹性伸缩或迁移。

针对 SDN/NFV 环境下动态 SFC 部署和重调整问题, 文献 [84] 提出了基于双深度 Q 网络的 DDQN-VNFPA 算法。该算法将 VNF 放置问题建模为二进制整数规划问题, 通过最小化包含 VNF 放置成本、运行成本和拒绝惩罚的加权成本函数进行优化。DDQN-VNFPA 通过离线训练与在线运行相结合的方式, 能够根据网络状态和预测的 SFC 请求动态优化 VNF 放置策略。文献 [85] 提出基于 DQN 和 M 最短路径算法的动态 SFC 部署和重调整方法。该方法采用双智能体协同工作机制, 一个负责动态部署 SFC, 另一个根据当前网络条件和 SFC 请求信息通过 VNF 迁移对底层网络进行重调整, 以提高服务请求接受率。

针对卫星网络中 SFC 迁移面临的资源受限、网络拓扑动态变化和传输时延等挑战, 文献 [86] 提出了一种基于深度强化学习的时延敏感型 SFC 迁移算法。该方案主要从 3 个方面进行优化: 首先, 在目标节点选择过程中引入资源评估机制, 优先选择资源利用率较低的节点; 其次, 设计自适应机制来感知卫星轨道和天气条件变化, 实现迁移策略的动态调整, 以应对网络环境的动态性; 最后, 针对实时性要求高的应用场景, 将时延最小化作为 SFC 迁移算法的优化目标, 通过优化路径选择来降低时延。该研究通过上述措施提高了 SFC 迁移算法在卫星通信环境下的适应性、智能性和性能表现, 为卫星网络中的服务提供了更可靠和高效的支持。

在移动边缘计算和 B5G/6G 网络环境下, 用户移动性导致的服务质量波动是 SFC 面临的重要挑战。移动性感知的 VNF 迁移通过预测用户移动轨迹、评估网络状态并进行及时迁移决策, 以维持服务质量并降低资源开销。文献 [87] 针对用户移动场景下的 SFC 迁移时机决策问题, 提出基于深度 Q 网络的 SFC 迁移时机决策算法。考虑到用户移动路径已知且到达时间可预测的情况, 该研究建立了用户到达时间预测模型, 将 SFC 迁移过程建模为马尔可夫决策过程, 通过智能体感知网络资源状态变化和 SFC 迁移信息, 选择合适的迁移时机, 以降低服务中断时间并提高迁移成功率。文献 [88] 针对多接入边缘计算环境中用户移动性导致的 SFC 动态迁移问题, 提出了 DRL 的 SFC 自适应动态迁移优化算法。该算法利用 DQN 感知网络资源状态和迁移需求, 通过对多个网络功能迁移路径的智能决策和带宽自适应分配, 实现了 SFC 的无缝迁移。

4.2 基于预测的动态重配置

网络环境的动态性和不确定性对 SFC 重配置提

出了严峻挑战。通过准确预测网络状态和资源需求变化,并据此及时调整部署策略,成为提升系统性能的重要研究方向。

文献 [89] 针对物联网场景下 SFC 动态重配置问题,提出了一种基于深度 Dyna-Q 的解决方案。该方案重点关注边缘云环境下,由于物联网业务流量动态变化导致的 VNF 实例资源需求波动问题。不同于传统静态部署机制,该方案能够在运行时动态调整 SFC 模式,通过预测 VNF 资源需求来优化 VNF 实例迁移和服务路径重路由配置。文献 [90] 提出了一种 SFC 主动重配置机制 SARM 算法来解决网络状态动态变化引起的服务质量下降问题。该方法将计算负载时间序列分解为长期趋势和短期波动 2 部分,分别采用长短期记忆网络 (long short-term memory, LSTM) 和 DQN 进行预测;同时假设 SFC 行为具有马尔可夫性质,利用隐马尔可夫模型 (hidden Markov model, HMM) 建模来预测未来资源需求。在此基础上, SARM 综合考虑端到端时延和重配置成本等多个因素,通过启发式算法平衡迁移和带宽成本,实现 SFC 的动态重配置。

4.3 可靠性导向的重配置

可靠性导向的重配置研究主要从故障容错、服务质量保障和优先级感知等多维度展开,旨在提升系统的鲁棒性和服务连续性。

针对网络拥塞和节点故障等导致的服务中断问题,研究者提出了多种基于优先级的容错方案。文

献 [91] 提出了优先级感知的 VNF 迁移方法 PAVM。该方法根据业务优先级设计差异化的迁移策略,通过深度强化学习优化目标节点选择,有效平衡了系统可靠性与资源利用效率。为进一步提升系统对动态负载的适应能力,文献 [92] 提出了一种基于 DRL 的自动缩放和优先级感知的 SFC 部署算法 APSD。该算法根据不同网络服务的 QoS 需求进行优先级划分,并基于 VNF 的弹性特征采用混合缩放策略,实现了对服务请求和工作负载动态变化的智能响应。

为提升系统的整体可靠性,研究者探索了基于分布式架构的故障恢复方案。文献 [93] 在 SDN/NFV 云环境中提出了面向容错的 SFC 优化方案,将容错导向的 SFC 弹性优化问题建模为 MDP。该方案通过设计包含能耗最小化、迁移成本最小化、收益最大化和负载均衡等因素的综合奖励函数,实现了多目标约束下的优化。特别地,文中提出的多智能体双深度 Q 网络方法将繁重的训练任务从控制平面转移到数据平面,有效降低了单点故障风险。文献 [94] 针对边缘网络中 SFC 迁移的可靠性问题,提出了基于多智能体深度强化学习的可靠性保障 SFC 迁移算法。该算法通过引入优势函数评估智能体性能,并采用中央注意力机制增强智能体间的协作。同时,基于迁移目标节点和链路容量设计了可靠性惩罚机制,在实时性、资源利用率和可靠性等方面均优于现有方法,为边缘网络中的可靠性保障提供了新思路。

基于以上分析,表 6 总结了近年来基于强化学习

Table 6 Comparison of SFC Reconfiguration Algorithms Based on Reinforcement Learning

表 6 基于强化学习的 SFC 重配置算法对比

重配置算法	强化学习算法	优化目标	VNF 重配置动作	应用场景
文献 [82]	DQN	QoS、能耗	迁移决策 (是否迁移、目标节点) 与资源重分配	5G 网络切片架构下 VNF 迁移优化
ISFCDA ^[83]	Double DQN	迁移成本、实例化成本、资源均衡度	重配置策略 (纵向扩展、横向扩展、迁移)	NFV 数据中心网络应对 SFC 动态需求的重配置
DDQN-VNFA ^[84]	Double DQN	部署成本、接受率、运行成本	对节点是否进行优化决策 (额外算法)	SDN/NFV 环境下应对动态负载的 VNF 重配置
MQDR ^[85]	DQN	接受率	二进制节点迁移决策	SDN/NFV 环境下处理动态到达 SFC 请求的重配置
LS-SFCM ^[86]	DDPG	时延、迁移成本	可迁移节点序号	卫星通信网络针对时延敏感型业务的 SFC 迁移优化
DQN-MTD ^[87]	DQN	宕机时间、迁移成功率	二进制节点迁移决策	边缘计算环境用户移动路径可预测场景的迁移时机优化
DRL-ADMO ^[88]	Double DQN	时延、迁移成功率、迁移时间	迁移路径选择	多接入边缘计算环境下用户移动导致的 SFC 动态迁移优化
DDQ ^[89]	Dyna-Q	收入、重配置成本	迁移与重路由决策	IoT 网络基于 GNN 预测资源需求的 SFC 重配置优化
SARM ^[90]	DQN	时延、迁移成本、带宽成本	下一个时隙的短期状态值	基于 LSTM 与 DQN 预测计算负载的主动重配置优化
PAVM ^[91]	DQN	时延、节点与链路负载率	迁移节点序号	NFV 环境下基于优先级感知的可靠 VNF 迁移优化
APSD ^[92]	DQN	时延、接受率、部署成本	重配置节点序号	NFV 环境下优先级感知与自动伸缩的重配置
SADDQN/MADDQN ^[93]	Double DQN / MADDQN	能耗、迁移成本、收入、负载均衡	迁移决策 (是否迁移、目标节点)	SDN/NFV 云环境下面向容错 SFC 弹性重配置
RA-SFCM ^[94]	A2C	可靠性、迁移成本	可迁移节点序号	边缘计算环境下多智能体的可靠性保障型 SFC 迁移优化

的 VNF 重配置算法特征。

从表 6 中可以看出, DQN 及其变体是当前 VNF 重配置领域最主流的强化学习方法, 其他先进算法如 PPO 等尚未在该领域得到充分应用。在优化目标方面, 大多数算法都采用多目标优化框架, 主要关注时延、迁移成本、能耗和请求接受率等性能指标, 反映了 VNF 重配置问题在保障服务质量与控制资源开销间的平衡需求。从重配置动作的设计特点来看, 现有研究呈现出明显的分化趋势: 一类算法(如 ISFCDA^[83], DDQN-VNFPA^[84])侧重于决策层面, 关注是否迁移及重配置策略的选择; 另一类算法(如 LS-SFCM^[86], PAVM^[91])则专注于执行层面, 主要确定具体的迁移节点序号。这种动作空间设计的差异性反映了不同应用场景下对 VNF 重配置精细化控制的不同需求。

5 面临的挑战与机遇

综合上述国内外研究现状, 当前基于强化学习的 SFC 部署研究虽取得了显著进展, 但仍面临诸多挑战。

5.1 算法设计层面

当前基于强化学习的 SFC 部署算法在泛化能力、训练效率和多任务协同等方面存在明显不足。首先, 当前基于强化学习的 SFC 部署模型过于依赖特定网络拓扑结构, 缺乏拓扑泛化能力。在大多数研究中, 状态空间设计局限于静态特征提取, 模型的设计和训练是在特定网络拓扑下进行的, 模型的优化目标和策略都与该拓扑高度耦合。一旦网络拓扑发生变化, 模型通常需要重新训练, 以适应新的拓扑环境。然而, 模型在拓扑泛化能力方面的不足, 严重制约了其在实际网络中的应用扩展性, 尤其是在动态网络拓扑频繁变化或跨网络环境应用的场景中。未来的研究需要探讨如何提升强化学习模型的拓扑无关性, 使得模型能够在不同拓扑结构间迁移, 并减少重新训练的成本。

在训练效率和收敛性方面, 强化学习算法在处理大规模网络的 SFC 部署问题时往往面临训练周期长、收敛性差等挑战。特别是在多智能体协同场景下, 环境动态从单个智能体的角度表现出非平稳性, 这导致传统的经验回放机制难以直接应用。同时, 智能体之间的相互影响进一步增加了训练的复杂性。在复杂的部署场景中, 有效的奖励信号往往稀疏且延迟, 智能体难以准确评估其行为的长期影响, 容易导致策略震荡或陷入局部最优。为应对这些挑战,

未来研究可从设计针对非平稳环境的自适应经验回放机制、构建分层式的奖励函数和值函数估计器促进稀疏奖励信号向密集反馈的转化、开发稳定的多智能体训练算法方向着手, 通过算法创新提升收敛性。

现有研究对 SFC 部署放置、调度、重配置各阶段缺乏系统性的协同优化机制。具体表现为在问题建模中对各阶段任务进行割裂式处理, 其优化目标设计相互孤立。这种分离处理方式忽视了各阶段决策行为之间的密切关联, 难以实现全局最优。未来研究应着重构建统一的端到端优化框架, 设计多层次的奖励机制以平衡短期与长期优化目标, 并探索基于因果推理的决策方法, 从而实现 SFC 部署各阶段的有效协同。

智能体协作与隐私保护的平衡成为重要研究方向。在多智能体协同决策场景中, 智能体间的信息共享与隐私保护构成了一对矛盾。过度共享可能导致隐私泄露, 而信息封闭又会影响决策效果。联邦强化学习为解决这一问题提供了新思路: 通过模型参数而非原始数据共享保护隐私; 利用联邦学习框架实现分布式协同优化; 结合差分隐私等技术增强安全性。

5.2 应用场景层面

在应用场景层面, 新型网络架构下的 SFC 部署面临诸多挑战。随着 5G/6G、算力网络等新型网络架构的快速演进, SFC 部署面临更加动态和复杂的应用环境。在车联网场景中, 高速移动的车辆节点频繁改变其地理位置和连接状态, 导致网络拓扑迅速变化。这种动态特性要求 SFC 部署具有更高的实时性, 能够快速调整服务链的位置与路径, 同时保证通信的可靠性。在空天地一体化网络中, 由于卫星节点的轨道运动和动态网络拓扑变化, SFC 部署需应对频繁的链路重构和跨域资源协调问题。同时, 大尺度传播时延带来的网络时延和抖动可能显著影响服务质量, 要求 SFC 部署算法具备精确的时延预测和优化能力。在算力网络中, 随着边缘计算、云计算和分布式计算资源的融合, 网络中不仅需要高效传递数据, 还需要协同调度分布式算力资源。算力网络中的 SFC 部署需要综合考虑计算资源和网络资源的联合优化。这些新型网络架构的特点要求 SFC 部署方法具备更强的环境适应能力和决策智能性。

此外, 目前针对并行 SFC 部署的研究相对较少。与串行 SFC 相比, 并行 SFC 不仅需要考虑 VNF 之间的依赖关系, 还需要根据底层网络状态动态调整并行化程度。特别是在复杂网络环境中, 如何将并行

SFC 的构建与部署调度、动态重配置等问题进行统一优化,仍然是一个亟待解决的挑战。这要求在现有研究基础上,进一步探索适应并行化特征的部署优化方法。

面对这些挑战与机遇,基于强化学习的 SFC 部署研究需在保持技术创新的同时,更加注重实际应用价值。通过算法优化、场景创新和技术融合,推动相关成果在实际网络中的落地应用,为未来网络的智能化发展提供有力支撑。

6 结 论

随着网络技术的快速演进,传统的 SFC 部署方法在应对动态网络环境和复杂部署场景时表现出诸多局限性,难以满足新型网络架构对实时性、可靠性和资源利用效率的严格要求。基于强化学习的 SFC 部署方法因其自适应性和智能化特点,成为近年来的研究热点。本文系统总结了基于强化学习的 SFC 部署研究现状,重点分析了强化学习在 SFC 放置、调度和重配置 3 个主要部署阶段的应用。通过分析现有研究在算法设计、性能优化和实际部署等方面的优势与局限性,探讨了该领域面临的挑战和未来研究方向。研究表明,基于强化学习的 SFC 部署方法相较于传统优化方法具有显著优势。强化学习通过与环境交互进行自适应学习,能够有效应对网络环境的动态变化。深度强化学习的引入解决了传统方法在处理高维状态空间时的局限性,多智能体框架的应用实现了分布式协同决策,提高了系统的可扩展性。然而,当前研究在强化学习模型的泛化能力、训练效率与收敛性、SFC 部署各阶段的协同优化机制以及新型网络架构下的应用探索等方面存在若干亟待解决的关键问题。未来研究应重点关注算法创新与实际应用的结合,在提升模型性能的同时,注重解决方案在实际网络环境中的部署落地。

作者贡献声明:爱因坦负责论文的整体设计,撰写并修改论文;阮宏伟提出指导意见并参与论文的修改完善;刘翰林负责部分资料的收集与整理工作;李华对论文选题、组织结构提供指导意见,并完成论文审阅工作。

参 考 文 献

- [1] International Telecommunication Union (ITU). ICT facts and figures 2024[R]. Geneva: International Telecommunication Union, 2024
- [2] CISCO. 2024 global networking trends report[R]. San Jose: Cisco Systems Inc, 2024
- [3] Mijumbi R, Serrat J, Gorricho J L, et al. Network function virtualization: State-of-the-art and research challenges[J]. *IEEE Communications Surveys & Tutorials*, 2015, 18(1): 236–262
- [4] Kreutz D, Ramos F M V, Verissimo P E, et al. Software-defined networking: A comprehensive survey[J]. *Proceedings of the IEEE*, 2014, 103(1): 14–76
- [5] Halpern J, Pignataro C. Service function chaining (SFC) architecture[R]. Fremont, CA: Internet Engineering Task Force, 2015
- [6] Bhamare D, Jain R, Samaka M, et al. A survey on service function chaining[J]. *Journal of Network and Computer Applications*, 2016, 75: 138–155
- [7] Xiao Yikai, Zhang Qixia, Liu Fangming, et al. NFVdeep: Adaptive online service function chain deployment with deep reinforcement learning[C]//Proc of the 27th Int Symp on Quality of Service. New York: ACM, 2019: 1–10
- [8] Kaelbling L P, Littman M L, Moore A W. Reinforcement learning: A survey[J]. *Journal of Artificial Intelligence Research*, 1996, 4: 237–285
- [9] Heo D N, Lange S, Kim H G, et al. Graph neural network based service function chaining for automatic network control[C]//Proc of the 21st Asia-Pacific Network Operations and Management Symp (APNOMS). Piscataway, NJ: IEEE, 2020: 7–12
- [10] Shah H A, Zhao Lian. Multiagent deep-reinforcement-learning-based virtual resource allocation through network function virtualization in Internet of things[J]. *IEEE Internet of Things Journal*, 2020, 8(5): 3410–3421
- [11] Cai Xuhui, Deng Hui, Deng Lingli, et al. Evolving NFV towards the next decade[R]. Sophia Antipolis, France: European Telecommunications Standards Institute, 2023
- [12] Zhang Biao, Fan Qilin, Zhang Xu, et al. A survey of VNF forwarding graph embedding in B5G/6G networks[J]. *Wireless Networks*, 2024, 30(5): 3735–3758
- [13] Zhou Weilin, Yang Yuan, Xu Mingwei. Network function virtualization technology research[J]. *Journal of Computer Research and Development*, 2018, 55(4): 675–688 (in Chinese)
(周伟林, 杨莹, 徐明伟. 网络功能虚拟化技术研究综述[J]. *计算机研究与发展*, 2018, 55(4): 675–688)
- [14] Kaur K, Mangat V, Kumar K. A comprehensive survey of service function chain provisioning approaches in SDN and NFV architecture[J]. *Computer Science Review*, 2020, 38: 100298
- [15] Schardong F, Nunes I, Schaeffer-Filho A. NFV resource allocation: A systematic review and taxonomy of VNF forwarding graph embedding[J]. *Computer Networks*, 2021, 185: 107726
- [16] Zu Jiachen, Hu Guyu, Yan Jiajie, et al. Resource management of service function chain in NFV enabled network: A survey[J]. *Journal of Computer Research and Development*, 2021, 58(1): 137–152 (in Chinese)
(祖家琛, 胡谷雨, 严佳洁, 等. 网络功能虚拟化下服务功能链的资源管理研究综述[J]. *计算机研究与发展*, 2021, 58(1): 137–152)
- [17] Yang Song, Li Fan, Trajanovski S, et al. Recent advances of resource

- allocation in network function virtualization[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2021, 32(2): 295–314
- [18] Santos G L, Bezerra D F, Rocha E S, et al. Service function chain placement in distributed scenarios: A systematic review[J]. *Journal of Network and Systems Management*, 2022, 30(1): 4
- [19] Pattaranantakul M, Vorakulpipat C, Takahashi T. Service function chaining security survey: Addressing security challenges and threats[J]. *Computer Networks*, 2023, 221: 109484
- [20] ETSI ISG NFV. Network functions virtualisation (NFV): Architectural framework[R]. Sophia Antipolis, France: European Telecommunications Standards Institute, 2013
- [21] John W, Pentikousis K, Agapiou G, et al. Research directions in network service chaining[C]//Proc of the 2013 IEEE SDN for Future Networks and Services (SDN4FNS). Piscataway, NJ: IEEE, 2013: 1–7
- [22] Bjorklund M. YANG—A data modeling language for the network configuration protocol (NETCONF)[R]. Fremont, CA: Internet Engineering Task Force, 2010
- [23] Mehrahadam S, Keller M, Karl H. Specifying and placing chains of virtual network functions[C]//Proc of the 3rd IEEE Int Conf on Cloud Networking (CloudNet). Piscataway, NJ: IEEE, 2014: 7–13
- [24] Binz T, Breitenbücher U, Kopp O, et al. TOSCA: Portable Automated Deployment and Management of Cloud Applications[M]//Advanced Web Services. Berlin: Springer, 2013: 527–549
- [25] Sun Chen, Bi Jun, Zheng Zhilong, et al. NFP: Enabling network function parallelism in NFV[C]//Proc of the Conf of the ACM Special Interest Group on Data Communication. New York: ACM, 2017: 43–56
- [26] Lin I C, Yeh Y H, Lin K C J. Toward optimal partial parallelization for service function chaining[J]. *IEEE/ACM Transactions on Networking*, 2021, 29(5): 2033–2044
- [27] Bari F, Chowdhury S R, Ahmed R, et al. Orchestrating virtualized network functions[J]. *IEEE Transactions on Network and Service Management*, 2016, 13(4): 725–739
- [28] Addis B, Belabed D, Bouet M, et al. Virtual network functions placement and routing optimization[C]//Proc of the 4th IEEE Int Conf on Cloud Networking (CloudNet). Piscataway, NJ: IEEE, 2015: 171–177
- [29] Liu Yi, Zhang Hongqi, Yang Yingjie, et al. A hierarchical method for survivable service function chain embedding[J]. *Journal of Computer Research and Development*, 2018, 55(4): 748–767 (in Chinese)
(刘艺, 张红旗, 杨英杰, 等. 一种区分等级的可生存服务功能链映射方法[J]. *计算机研究与发展*, 2018, 55(4): 748–767)
- [30] Riera J F, Hesselbach X, Escalona E, et al. On the complex scheduling formulation of virtual network functions over optical networks[C]//Proc of the 16th Int Conf on Transparent Optical Networks (ICTON). Piscataway, NJ: IEEE, 2014: 1–5
- [31] Qu Long, Assi C, Shaban K. Delay-aware scheduling and resource optimization with network function virtualization[J]. *IEEE Transactions on Communications*, 2016, 64(9): 3746–3758
- [32] Yang Yong, Meng Xiangru, Kang Qiaoyan, et al. Dynamic service function chain migration method based on resource requirements prediction[J]. *Journal of Computer Research and Development*, 2023, 60(5): 1151–1163 (in Chinese)
(阳勇, 孟相如, 康巧燕, 等. 基于资源需求预测的动态服务功能链迁移方法[J]. *计算机研究与发展*, 2023, 60(5): 1151–1163)
- [33] Eramo V, Miucci E, Ammar M, et al. An approach for service function chain routing and virtual function network instance migration in network function virtualization architectures[J]. *IEEE/ACM Transactions on Networking*, 2017, 25(4): 2008–2025
- [34] Liu Junjie, Lu Wei, Zhou Fen, et al. On dynamic service function chain deployment and readjustment[J]. *IEEE Transactions on Network and Service Management*, 2017, 14(3): 543–553
- [35] Bellman R. A Markovian decision process[J]. *Journal of Mathematics and Mechanics*, 1957, 6(5): 679–684
- [36] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. *Nature*, 2015, 518(7540): 529–533
- [37] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484–489
- [38] Liu Quan, Zhai Jianwei, Zhang Zongzhang, et al. A survey on deep reinforcement learning[J]. *Chinese Journal of Computers*, 2018, 41(1): 1–27 (in Chinese)
(刘全, 翟建伟, 章宗长, 等. 深度强化学习综述[J]. *计算机学报*, 2018, 41(1): 1–27)
- [39] Ding Shifei, Du Wei, Zhang Jian, et al. Research progress of multi-agent deep reinforcement learning[J]. *Chinese Journal of Computers*, 2024, 47(7): 1547–1567 (in Chinese)
(丁世飞, 杜威, 张健, 等. 多智能体深度强化学习研究进展[J]. *计算机学报*, 2024, 47(7): 1547–1567)
- [40] Kim S I, Kim H S. A research on dynamic service function chaining based on reinforcement learning using resource usage[C]//Proc of the 9th Int Conf on Ubiquitous and Future Networks (ICUFN). Piscataway, NJ: IEEE, 2017: 582–586
- [41] Sun Jian, Huang Guanhua, Sun Gang, et al. A Q-learning-based approach for deploying dynamic service function chains[J]. *Symmetry*, 2018, 10(11): 646
- [42] Watkins C J C H, Dayan P. Q-learning[J]. *Machine Learning*, 1992, 8(3): 279–292
- [43] Chen Jing, Chen Jia, Zhang Hongke. DRL-QOR: Deep reinforcement learning-based QoS/QoE-aware adaptive online orchestration in NFV-enabled networks[J]. *IEEE Transactions on Network and Service Management*, 2021, 18(2): 1758–1774
- [44] Nouruzi A, Zakeri A, Javan M R, et al. Online service provisioning in NFV-enabled networks using deep reinforcement learning[J]. *IEEE Transactions on Network and Service Management*, 2022, 19(3): 3276–3289
- [45] Li Guanglei, Feng Bohao, Zhou Huachun, et al. Adaptive service function chaining mappings in 5G using deep Q-learning[J]. *Computer Communications*, 2020, 152: 305–315
- [46] Liu Guangyuan, Chen Shiyang, Pang Ziyuan. Deployment mechanism for maximizing revenue from online service function chains in edge cloud environments[J]. *Journal of Tsinghua University: Science and Technology*, 2025, 65(8): 1516–1529
(刘光远, 陈世莹, 庞紫园. 边缘云环境中收益最大化的在线服务功能链部署机制[J]. *清华大学学报: 自然科学版*, 2025, 65(8): 1516–1529)

- [47] Quang P T A, Hadjadj-Aoul Y, Outtagarts A. A deep reinforcement learning approach for VNF forwarding graph embedding[J]. *IEEE Transactions on Network and Service Management*, 2019, 16(4): 1318–1331
- [48] Tang Lun, He Lanqin, Lian Qinyi, et al. Virtual network function placement optimization algorithm based on improve deep reinforcement learning[J]. *Journal of Electronics & Information Technology*, 2021, 43(6): 1724–1732 (in Chinese)
(唐伦, 贺兰钦, 连沁怡, 等. 基于改进深度强化学习的虚拟网络功能部署优化算法[J]. *电子与信息学报*, 2021, 43(6): 1724–1732)
- [49] He Nan, Yang Song, Li Fan, et al. A-DDPG: Attention mechanism-based deep reinforcement learning for NFV[C]//Proc of the 29th IEEE/ACM Int Symp on Quality of Service (IWQOS). Piscataway, NJ: IEEE, 2021: 1–10
- [50] He Nan, Yang Song, Li Fan, et al. Leveraging deep reinforcement learning with attention mechanism for virtual network function placement and routing[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2023, 34(4): 1186–1201
- [51] Liu Yicen, Zhang Junning. Service function chain embedding meets machine learning: Deep reinforcement learning approach[J]. *IEEE Transactions on Network and Service Management*, 2024, 21(3): 3465–3481
- [52] Huang Zhongwei, Zhong Weijian, Li Dagang, et al. Delay constrained SFC orchestration for edge intelligence-enabled IIoT: A DRL approach[J]. *Journal of Network and Systems Management*, 2023, 31(3): 53
- [53] Wang Lei, Mao Weixi, Zhao Jin, et al. DDQP: A double deep Q-learning approach to online fault-tolerant SFC placement[J]. *IEEE Transactions on Network and Service Management*, 2021, 18(1): 118–132
- [54] Qu Hua, Wang Ke, Zhao Jihong. Reliable service function chain deployment method based on deep reinforcement learning[J]. *Sensors*, 2021, 21(8): 2733
- [55] Santos G L, Lynn T, Kelner J, et al. Availability-aware and energy-aware dynamic SFC placement using reinforcement learning[J]. *The Journal of Supercomputing*, 2021, 77(11): 12711–12740
- [56] Santos G L, Endo P T, Lynn T, et al. A reinforcement learning-based approach for availability-aware service function chain placement in large-scale networks[J]. *Future Generation Computer Systems*, 2022, 136: 93–109
- [57] Marler R T, Arora J S. The weighted sum method for multi-objective optimization: new insights[J]. *Structural and Multidisciplinary Optimization*, 2010, 41(6): 853–862
- [58] Konak A, Coit D W, Smith A E. Multi-objective optimization using genetic algorithms: A tutorial[J]. *Reliability Engineering & System Safety*, 2006, 91(9): 992–1007
- [59] Bi Yu, Meixner C C, Bunyakitanon M, et al. Multi-objective deep reinforcement learning assisted service function chains placement[J]. *IEEE Transactions on Network and Service Management*, 2021, 18(4): 4134–4150
- [60] Li Huaping, Kordi M E. DSPPV: Dynamic service function chains placement with parallelized virtual network functions in mobile edge computing[J]. *Internet of Things*, 2023, 22: 100733
- [61] Jia Jing, Hua Jing. Dynamic SFC placement with parallelized VNFs in data center networks: A DRL-based approach[J]. *ICT Express*, 2024, 10(1): 104–110
- [62] Sun Penghao, Lan Julong, Li Junfei, et al. Combining deep reinforcement learning with graph neural networks for optimal VNF placement[J]. *IEEE Communications Letters*, 2020, 25(1): 176–180
- [63] Battaglia P W, Hamrick J B, Bapst V, et al. Relational inductive biases, deep learning, and graph networks[J]. arXiv preprint, arXiv: 1806.01261, 2018
- [64] Xie Yanghao, Huang Lin, Kong Yuyang, et al. Virtualized network function forwarding graph placing in SDN and NFV-enabled IoT networks: A graph neural network assisted deep reinforcement learning method[J]. *IEEE Transactions on Network and Service Management*, 2021, 19(1): 524–537
- [65] Qiu Rixuan, Bao Jiawen, Li Yuancheng, et al. Virtual network function deployment algorithm based on graph convolution deep reinforcement learning[J]. *The Journal of Supercomputing*, 2023, 79(6): 6849–6870
- [66] Fan Qilin, Pan Pan, Li Xiuhua, et al. DRL-D: Revenue-aware online service function chain deployment via deep reinforcement learning[J]. *IEEE Transactions on Network and Service Management*, 2022, 19(4): 4531–4545
- [67] Zhu Yuchao, Yao Haipeng, Mai Tianle, et al. Multiagent reinforcement-learning-aided service function chain deployment for Internet of things[J]. *IEEE Internet of Things Journal*, 2022, 9(17): 15674–15684
- [68] Xiao Da, Zhang J A, Liu Xin, et al. A two-stage GCN-based deep reinforcement learning framework for SFC embedding in multi-datacenter networks[J]. *IEEE Transactions on Network and Service Management*, 2023, 20(4): 4297–4312
- [69] Wang Shaoyang, Yuen C, Ni Wei, et al. Multiagent deep reinforcement learning for cost-and delay-sensitive virtual network function placement and routing[J]. *IEEE Transactions on Communications*, 2022, 70(8): 5208–5224
- [70] Wang Kan, Yuan Peng, Jan M A, et al. Digital twin-assisted service function chaining in multi-domain computing power networks with multi-agent reinforcement learning[J]. *Future Generation Computer Systems*, 2024, 158: 294–307
- [71] Wang Shuyi, Cao Haotong, Yang Longxiang, et al. GCN-based multi-agent deep reinforcement learning for dynamic service function chain deployment in IoT[J]. *IEEE Transactions on Consumer Electronics*, 2024, 70(3): 6105–6118
- [72] Wang Tianfeng, Zu Jiachen, Hu Guyu, et al. Adaptive service function chain scheduling in mobile edge computing via deep reinforcement learning[J]. *IEEE Access*, 2020, 8: 164922–164935
- [73] He Rui, Ren Bangbang, Xie Junjie, et al. A reinforcement learning method for scheduling service function chains with multi-resource constraints[J]. *Computer Networks*, 2023, 235: 109985
- [74] Yao Jiamin, Wang Junli, Wang Cheng, et al. DRL-based VNF cooperative scheduling framework with priority-weighted delay[J]. *IEEE Transactions on Mobile Computing*, 2024, 23(12): 11375–11388
- [75] Jia Ziye, Cao Yilu, He Lijun, et al. Service function chain dynamic scheduling in space-air-ground integrated networks[J]. *IEEE Transactions on Vehicular Technology*, 2025, 74(7): 11235–11248

- [76] Li Junling, Shi Weisen, Zhang Ning, et al. Delay-aware VNF scheduling: A reinforcement learning approach with variable action set[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 7(1): 304–318
- [77] Akbari M, Abedi M R, Joda R, et al. Age of information aware VNF scheduling in industrial IoT using deep reinforcement learning[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(8): 2487–2500
- [78] Kuai Zhenran, Wang Tianyu, Wang Shaowei. Fair virtual network function mapping and scheduling using proximal policy optimization[J]. *IEEE Transactions on Communications*, 2022, 70(11): 7434–7445
- [79] Zhang Qinghua, Zhang Xianchao, Wang Yin hao, et al. Service function chains scheduling of information networks used for emergency treatment[J]. *Acta Electronica Sinica*, 2023, 51(11): 3128–3136 (in Chinese)
(张庆华, 张先超, 王寅昊, 等. 面向医疗急救的信息网络服务功能链调度方法[J]. *电子学报*, 2023, 51(11): 3128–3136)
- [80] Lu Yuxi, Zhang Peiyong, Duan Youxiang, et al. Dynamic scheduling of iov edge cloud service functions under NFV: A multi-agent reinforcement learning approach[J]. *IEEE Transactions on Vehicular Technology*, 2023, 73(4): 5730–5741
- [81] Liu Zhiwei, Shu Zhaogang, Chen Shuwu, et al. Low-latency virtual network function scheduling algorithm based on deep reinforcement learning[J]. *Computer Networks*, 2024, 246: 110418
- [82] Tang Lun, Zhou Yu, Tan Qi, et al. Virtual network function migration algorithm based on reinforcement learning for 5G network slicing[J]. *Journal of Electronics & Information Technology*, 2020, 42(3): 669–677 (in Chinese)
(唐伦, 周钰, 谭颀, 等. 基于强化学习的 5G 网络切片虚拟网络功能迁移算法[J]. *电子与信息学报*, 2020, 42(3): 669–677)
- [83] Wang Yuantao, Shu Zhaogang, Chen Shuwu, et al. A cost and demand sensitive adjustment algorithm for service function chain in data center network[J]. *Computer Networks*, 2024, 242: 110254
- [84] Pei Jianing, Hong Peilin, Pan Miao, et al. Optimal VNF placement via deep reinforcement learning in SDN/NFV-enabled networks[J]. *IEEE Journal on Selected Areas in Communications*, 2019, 38(2): 263–278
- [85] Ran Jing, Wang Wenkai, Hu Hefei. Dynamic service function chain deployment and readjustment method based on deep reinforcement learning[J]. *Sensors*, 2023, 23(6): 3054
- [86] Zhang Peiyong, Li Yilin, Tan Lizhuang, et al. Latency-sensitive service function chains intelligent migration in satellite communication driven by deep reinforcement learning[J]. *Transactions on Emerging Telecommunications Technologies*, 2024, 35(11): e70006
- [87] Hu Hefei, Zhang Wei, Xu Lingyi, et al. A mobility-aware service function chain migration strategy based on deep reinforcement learning[J]. *Journal of Network and Systems Management*, 2023, 31(1): 21
- [88] Xu Lingyi, Liu Wenbin, Wang Zhiwei, et al. Mobile-aware service function chain intelligent seamless migration in multi-access edge computing[J]. *Journal of Network and Systems Management*, 2024, 32(3): 49
- [89] Liu Yicen, Lu Yu, Li Xi, et al. On dynamic service function chain reconfiguration in IoT networks[J]. *IEEE Internet of Things Journal*, 2020, 7(11): 10969–10984
- [90] Cai Jun, Qian Kaili, Luo Jianzhen, et al. SARM: Service function chain active reconfiguration mechanism based on load and demand prediction[J]. *International Journal of Intelligent Systems*, 2022, 37(9): 6388–6414
- [91] Qu Hua, Wang Ke, Zhao Jihong. Priority-awareness VNF migration method based on deep reinforcement learning[J]. *Computer Networks*, 2022, 208: 108866
- [92] Yu Xue, Wang Ran, Hao Jie, et al. Priority-aware deployment of autoscaling service function chains based on deep reinforcement learning[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(3): 1050–1062
- [93] Chen Jing, Chen Jia, Guo Kuo, et al. Fault tolerance oriented SFC optimization in SDN/NFV-enabled cloud environment based on deep reinforcement learning[J]. *IEEE Transactions on Cloud Computing*, 2024, 12(1): 200–218
- [94] Li Yilin, Zhang Peiyong, Kumar N, et al. Reliability-assured service function chain migration strategy in edge networks using deep reinforcement learning[J]. *Journal of Network and Computer Applications*, 2024, 231: 103999



Ai Yintan, born in 1994. PhD candidate. His main research interests include software-defined networking, service function chaining, and multi-agent reinforcement learning.

爱因坦, 1994年生。博士研究生。主要研究方向为软件定义网络、服务功能链、多智能体强化学习。



Ruan Hongwei, born in 1981. PhD. His main research interests include software-defined networking and software testing. (csrhw@imu.edu.cn)

阮宏玮, 1981年生。博士。主要研究方向为软件定义网络、软件测试。



Liu Hanlin, born in 1996. PhD candidate. Her main research interests include intent-based networking and graph neural networks. (imucslhl@mail.imu.edu.cn)

刘翰林, 1996年生。博士研究生。主要研究方向为意图驱动网络、图神经网络。



Li Hua, born in 1964. PhD, professor, PhD supervisor. Her main research interests include cloud-network integration, big data analysis and evaluation methods, and software service computing and testing.

李华, 1964年生。博士, 教授, 博士生导师。主要研究方向为云网融合、大数据分析 with 评估方法、软件服务计算与测试。