

知识定义多模态网络按需服务体系研究

王敬宇, 庄子睿

(北京邮电大学网络与交换技术国家重点实验室, 北京 100876)

摘要: 针对管理、控制、数据三平面解耦的未来移动通信网络系统, 提出了一种知识定义多模态网络按需服务体系架构。该架构仿照生物多态性的原理, 将“网络知识”作为贯穿多个平面的“基因”主干, 分场景、分层次地提取关键局部网络知识。通过构建逻辑统一的网络知识空间图谱, 根据具体业务对特定知识的依赖诉求, 不同的局部知识之间可以进行交换并形成有机联动。网络知识将能够面向不同服务场景、跨越多级服务层次、融合多种服务指标, 为网络整体的优化管理提供引导和支撑。知识定义多模态网络可以帮助移动通信网络应对复杂多变的业务需求, 使最终用户动态而多样化的需求可以得到及时、有效的满足和保障。

关键词: 知识定义网络; 多模态网络; 服务体验; 网络智能化; 人工智能

中图分类号: TN91/TN92

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2022076

Research on a knowledge-defined polymorphic network attainable service architecture

WANG Jingyu, ZHUANG Zirui

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: Given the consideration of future generation mobile communication network systems with decoupled administration, control, and data planes, a knowledge-defined polymorphic network (KDPN) attainable service architecture was proposed. In the proposed architecture, while mimicking the biological polymorphism, “network knowledge” was acted as a genetic core connecting all the planes, where local network knowledge was extracted from different scenarios and layers of the network. By constructing a network knowledge graph with unified logic, and depending on the service requirement for specific knowledge, different local knowledge could be exchanged and collaborated. Network knowledge will provide guidance and foundation for general network optimization and management across different service scenarios, service layers, and service objectives. KDPN will help the communication networks cope with sophisticated and dynamic service demands and provide dynamic and diversified services to general users promptly and effectively with quality of experience assurances.

Keywords: knowledge defined networking, polymorphic network, quality of experience, network intelligence, artificial intelligence

0 引言

随着互联网到物联网进再到万物互联的发展, 信息的传递也从人的互联变成物理设备的互联, 进

而变成人、物、数据的广泛互联。接入网络的信息端点越来越多, 导致信息传递的频率不断增加, 对信息时效性的要求也越来越高。这些变化导致作为信息传输基础设施的网络需要不断地提升与优化

收稿日期: 2022-01-05; 修回日期: 2022-03-21

通信作者: 庄子睿, zhuangzirui@bupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62071067, No.62101064); 北京邮电大学-中国移动研究院联合创新中心基金资助项目

Foundation Items: The National Natural Science Foundation of China (No.62071067, No.62101064), Beijing University of Posts and Telecommunications-China Mobile Research Institute Joint Innovation Center

性能。与此同时,随着后摩尔时代的到来,计算能力的提升开始变慢,计算资源的获得越来越昂贵,网络服务提供商不能再以不断增加计算硬件的方式来提升关键节点的处理容量。

在应用层面上,则涌现出各类日新月异的智能服务与产业业务,比如以虚拟现实、增强现实为代表的智能服务;以智慧交通、泛在物联为代表的智能城市服务;以自动驾驶决策控制为代表的智能计算服务;以智慧家居、绿色办公为代表的智能楼宇服务。下一代通信系统将承载、支撑更加个性化、泛在化、情景化的诸多服务,网络中将涌现大规模自主接入和智能决策节点,产生大量算力与数据资产。这些服务需要在贴近用户的边缘侧提供超高容量的智能决策能力,有着高可靠、低时延、快速响应的服务质量要求。智能决策能力高度依赖于计算资源,但是在现有成熟的方案中,计算资源往往更多地集中于云端数据中心,并非大量存在于最终用户周围。

边缘计算则是尝试解决这一问题的很好的思路^[1-13]。边缘计算在云端的智能业务逻辑和终端的用户设备之间部署服务计算资源,可以就近完成对用户数据的处理,响应用户的服务请求。从用户体验的角度上看,通过不断地向用户侧进行推进,边缘计算可以为最终用户提供融合而便捷的计算、存储与通信资源,有效地降低服务响应时间,提升服务质量。从智能业务部署的角度上看,应用了边缘计算框架之后,原本运行于云端数据中心的智能业务可以将一部分的计算任务卸载到边缘侧,减小了云端数据中心和用户设备之间的直接数据交换量,节省了网络带宽,能够在一定程度上减缓主干网络的压力,减小智能业务的服务容量预留成本。从网络运营商的角度上看,边缘计算模式可以对局域的计算与存储资源加以利用,使用各类多级缓存机制,用存储能力换取通信成本,减小在传输数据与内容过程中的能源消耗,因而边缘计算的引入也是后 5G 和未来 6G 的重要发展趋势之一。

随着边缘计算和各类业务相辅相成不断演进,越来越多的边缘设备接入通信网络中,既包含服务侧的资源提供设备,也包含用户侧的资源消耗设备,下一代通信网络将呈现超密集组网、全域感知、多形态共存等特点^[7,9,14-16]。通信系统将逐渐解耦并与各类网络融合,尤其是边缘网络将处于高度开放、异构、动态的环境中。

超大规模的边缘设备将提供前所未有的计算、感知和通信资源,但是其组织形式的极端复杂性与运行状态高动态的不稳定性对通信网络的组织和管理提出了挑战,需要突破传统网络二、三层分立的概念,跨越网络层次对路由与调度进行优化控制。如何寻找优秀的边缘网络智能路由与流量调度方案,以面对不断增加的以及差异化的网络服务体验要求,成为亟待解决的重要问题。

与此同时,网络服务的组织关系也在发生着变化。服务供应商依赖于特定的服务平台来提供服务^[17]。传统上看,服务供应商和服务平台的关系是相对简单的:单一供应商对应单一服务平台,比如特定的移动终端操作系统和桌面操作系统;多个供应商共同运作于同一个服务平台,比如互联网。随着服务的不断发展和各种新服务的涌现,服务供应商和服务平台的关系也变得更加复杂化和多层次化^[18],比如服务供应商会针对不同的操作系统开发应用软件并提供服务,逐步产生多供应商对应多服务平台的对应关系。随着网络虚拟化的演进,网络侧自身也开始呈现多平台化的复杂层次关系:多个网络供应商提供基础网络设施和资源,并由多个分销商共享、复用这些资源并提供给最终用户;服务提供商也会通过租用出口带宽或者通过内容分发网络(CDN, content distribution network)等途径跨越多个网络分销商来为最终用户提供服务。尽管每个平台内部都会尝试使用流量工程(TE, traffic engineering)等方法来优化网络流量的分布,但是由于端口、链路等基础网络设施和资源共享化的特性,更优化的网络资源管理和服务体验保障难以仅依靠单个平台的内部信息来实现^[19]。

在此背景下,为了达到极致性能服务体验要求,需要网络管控对服务需求与网络交付进行精准拟合,提供网络资源调配的快速优化、决策与控制。将感知、计算和通信融合,使用机器学习(ML, machine learning)和人工智能(AI, artificial intelligence)等技术将智能植入通信网络来进一步优化通信系统的资源分配、提升通信系统的服务质量是学术界和产业界的共识^[20-26]。然而,受服务增长的非均匀性、分散资源的低复用性、网络行为的复杂性和不确定性等因素的共同影响和相互制约,即便是使用智能化的网络节点,其朴素组合难以应对随需而变、高动态性的网络特征,难以发挥出智能节点的优势,无法实现通信系统资源的随愿共享。

受网络服务的多样性、网络设备的异构性和网络资源分布的差异性等因素的影响，通信网络体系需要针对各个方面所体现出的多态性加以理解、适配、管理和控制。比如多态控制的概念将基础光网络进行切片以支持不同种类的服务^[27]，针对不同的服务在每个切片中部署专用的网络资源和特定的交换设备。Chen 等^[28]设计了一种多态化协议以实现在移动自组织网络（MANET, mobile ad-hoc network）下功率、移动性和密度感知的路由组织。面对多样化的服务，多态路由（PR, polymorphic routing）^[29]根据不同业务和网络状态的特征，从基态派生出特定的路由方案。多模态智慧网络^[30-32]跨越网络的多个运行层次和网络功能的多个维度，提出了一种多态化的网络结构建模方法，用以提升网络能力、促进网络的进一步演化。

通信网络存在多态性，自然界也存在着多态性：比如在神经网络中，同一组神经元可以表现出不同的行为和功能^[33]；在生态系统中，同一种群中可以存在多种表型^[28]；在遗传物质的级别上，同一基因型可以表达出不同的性状^[34]。在上述例子中，多态性带来了 2 个重要优势：一方面，同一组信息可以表现出不同的行为、功能和模式来应对不同的环境状态，大大地提高了信息的利用效率；另一方面，优势信息也可以在代际之间进行传播，以便进行持续性的优化。类似地，在通信网络系统的体系结构设计中

可以借鉴该思路：一方面，使用网络结构的建模方法应当能够支撑不断增加、更新、变化的众多服务，以应对差异化的网络状态和网络环境变化；另一方面，用来帮助优化、决策的关键信息应当能够在不同的网络之间或者是在不同的智能节点之间进行传递，提升网络整体的优化和决策效率。

针对上述两点要求，本文认为，知识定义的多模态网络将是未来网络体系构建的一条重要线索。一方面，多态化的建模和设计可以有效地统一不同网络服务、不同网络层次、不同网络环境的差异性；另一方面，知识定义的方法可以在不同的管理位置、管理目标之间共享有价值的键信息，促进网络知识的复用。如图 1 所示，首先，知识定义多模态网络将能够回答“是什么”的描述性问题：网络有哪些参与者、具有什么不同的资源、网络的行为和变化趋势等。其次，知识定义多模态网络将能够回答“为什么”的理解性问题：哪些网络节点之间存在连接关系、哪些网络服务之间存在相互依赖、用户如何从网络中获取服务、不同的用户行为会对网络产生怎样的影响。最后，知识定义多模态网络将能够回答“怎么办”的执行性问题：为了实现特定的网络优化目标，应当采取何种优化、仿真、决策或是搜索算法。

针对多样化的网络服务需求，本文提出了一种在具有异构设备组成和差异化的资源分布条件下，

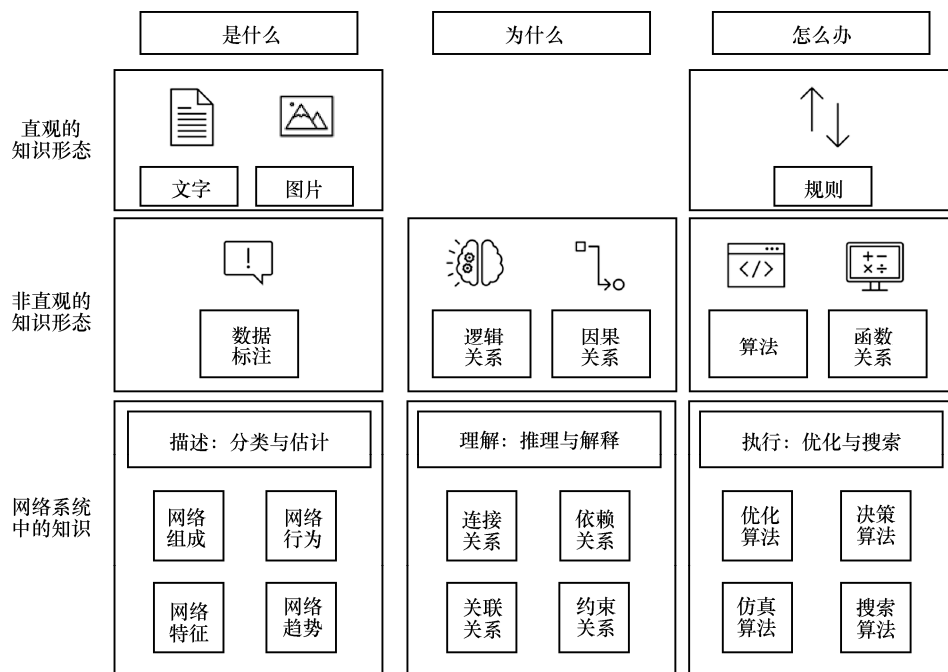


图 1 网络系统中的知识划分

共享优化决策信息、在线持续优化的知识定义多模态网络按需服务体系,支持对各种网络服务动态需求的满足,快速、协作、持续性地优化网络资源调配。本文进一步展示了知识定义多模态网络的建模、优化和部署方式,以消息传递图神经网络实现了知识交互过程。最终,以求解异构场景下的流量工程问题为应用案例,验证了本文方案的实际效果,如对域内和域间流量的收敛速度都体现出明显优势。

1 相关背景

1.1 多模态网络

Qiao 等^[27]使用多态控制的概念将基础光网络进行切片以支持不同类型的服务,针对不同的服务在每个切片中部署专用的网络资源和特定的交换设备。Chen 等^[28]设计了一种多态化协议以实现在 MANET 下功率、移动性和密度感知的路由组织。随着网络组成的异构性和网络服务的多样性的进一步加强,近年来有更多的研究聚焦在多模态网络上。比如多态路由^[29]根据不同业务和网络状态的特征,从基态派生出特定的路由方案,使网络路由能够动态地适应多样化的业务服务。全维可定义的多模态智慧网络^[30-32]从网络体系结构设计的高度,跨越网络的多个运行层次和网络功能的多个维度,提出了一种多态化的网络结构建模方法,用以提升网络能力、促进网络的进一步演化。

1.2 知识定义网络

2003 年,Clark 等^[35]提出了知识平面(KP, knowledge plane)这一概念,试图将智能植入网络的基础架构中去,通过知识来减少网络设计、部署、运营和维护的成本。2017 年,Mestres 等^[36]进一步提出知识定义网络(KDN, knowledge-defined networking),将软件定义网络和人工智能技术结合起来,进一步实现知识平面中提出的内生智能网络。知识加数据驱动的学习模型被认为是实现未来网络的智能管控方法的基础^[37]。

Hyun 等^[38]基于知识定义网络使用带内网络测量来实现自驱动的闭环网络管理。Pham 等^[39]使用深度强化学习实现了一种服务质量感知的网络路由方法。Xie 等^[40]回顾了软件定义网络中应用机器学习技术的一系列工作,并总结了分布式控制器的联合优化控制是亟待探索的研究方向。

1.3 网络智能化管理

飞速的技术和市场变化促使着软件化的网络组件和愈加强烈的网络智能化需求^[41]。从架构上看,CogNet 构建了统一化的认知知识网络,以持续、分布式地存储和利用时空域网络经验^[42]。从智能化的节点上看,分布式的多智能体系统是网络智能化的一个重要的发展方向^[43]。人工智能技术是未来网络智能化的使能技术^[44],而深度强化学习则是最突出的技术方向之一,在智能车联网络的移动计算卸载^[45]、用户需求分布变化情况下的网络路由优化^[46]、多资源联合优化^[47]等方向的应用上都看到了显著的成效。

2 知识定义多模态网络按需服务体系

知识定义网络参照软件定义网络解耦的数据平面和控制平面,并在上层添加了知识平面,一方面可以灵活地获取网络的全局视角状态信息,另一方面可以使用人工智能技术对这些信息进行处理,抽取有益于网络管理的知识。然而,网络本身具有异构特性,不同场景、不同层次、不同位置的网络可能具有差异化的可管理性;特别是在考虑面向用户的业务网络时,多样化的用户需求的突发性、动态性、不平衡性也给网络带来了差异化的网络优化和管理目标。而在已有的知识定义网络中,尽管研究者提出了使用人工智能技术来构建知识平面,但是并没有从网络架构的高度仔细探讨如何应对这种差异化的特性。与此同时,多模态网络架构所提出的多态化的网络结构建模方法,恰恰可以跨越网络的多个运行层次和网络功能的多个维度,是面对网络异构特性和优化目标差异化特性下行之有效的建模思想。

因此,在上述已有架构的基础上,本文进一步提出了知识定义多模态网络按需服务体系架构,使用管理、控制、数据解耦的三平面划分,遵循多模态网络的网络结构建模思想,使用多态化的方法对差异化的网络知识进行建模、求解和应用。知识定义多模态网络的最终目标是在不同的网络场景和层次中定义并学习其网络知识,用以自适应地提升网络能力、促进网络的自我演化。按需服务的知识定义多模态网络逻辑上可分为 3 个层次:数据平面、控制平面和知识平面,如图 2 所示。控制平面作为连接知识平面和数据平面的桥梁,向下从数据平面上测量和采集网络的原始

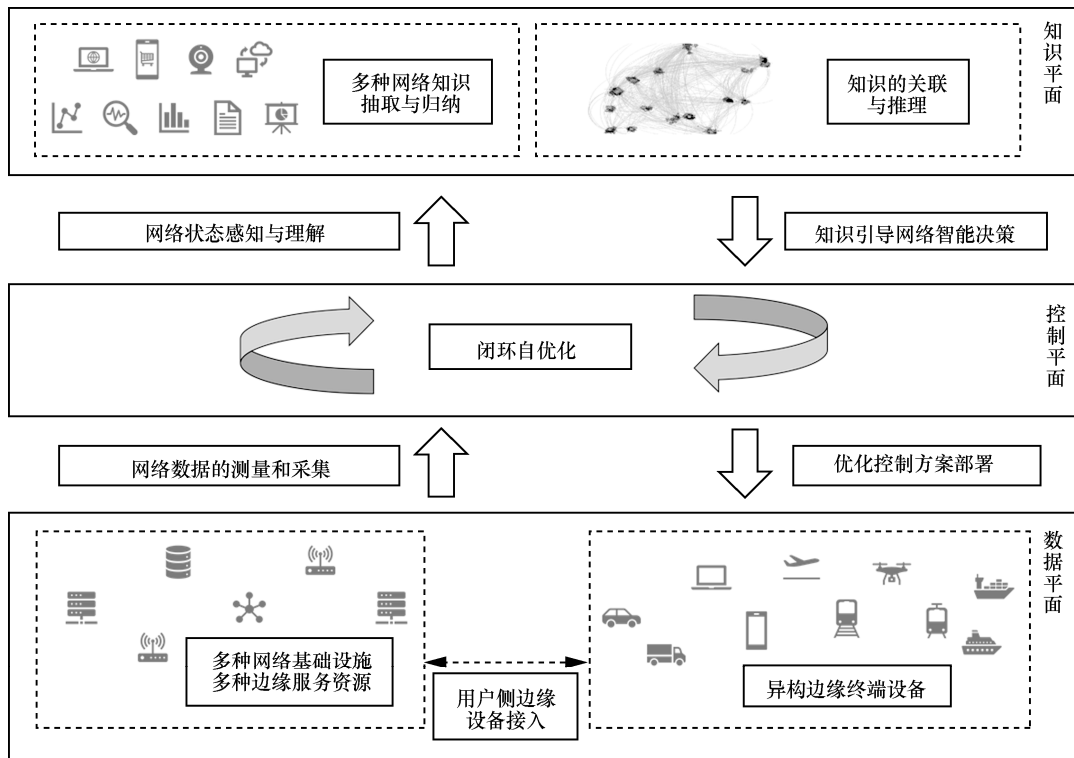


图 2 知识定义多模态网络

状态数据信息，并部署网络的优化控制方案；向上提供数据支撑，以供知识平面对网络状态数据进行感知和理解。知识平面是知识定义多模态网络的一个重要环节，一方面需要能够抽取、归纳多种网络知识信息，另一方面还要能够对多种网络知识进行有机的组合、发现网络知识之间的关联关系，并利用网络知识图进行网络优化管理的推理和决策、支撑网络的智能化管理。

知识定义多模态网络超越已有的知识定义网

络，在多个维度上引入了多态化的建模结构，可以更加精准、细粒度地支撑来自用户的动态、多样的服务请求。按需服务的知识定义多模态网络的实现路径如图 3 所示。具体地，涉及如下几个方面。

1) 服务场景的多态化。现有的 5G 网络提出了三大典型场景，即增强型移动宽带（eMBB, enhanced mobile broadband）、超可靠低时延通信（URLLC, ultra-reliable and low-latency communication）和大连接物联网（mMTC, massive machine-type

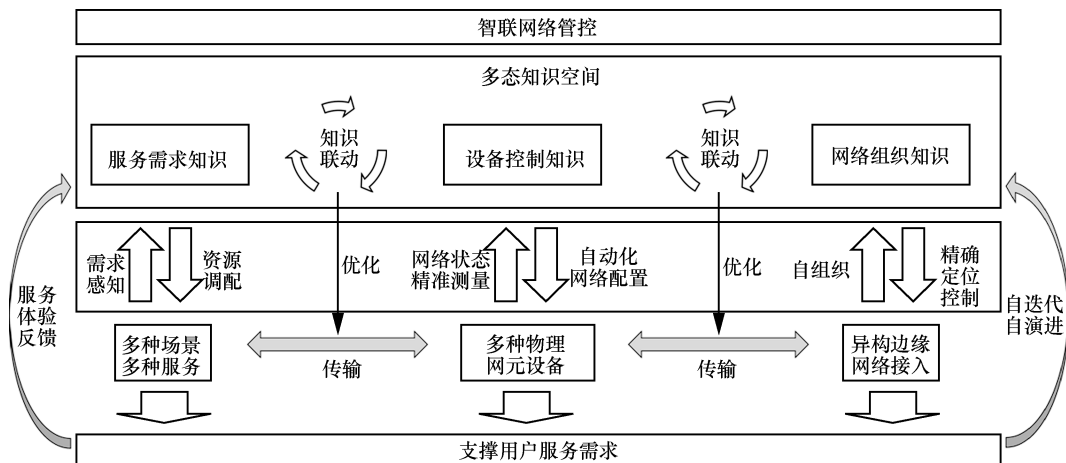


图 3 按需服务的知识定义多模态网络的实现路径

communication), 分别对应吞吐量、时延和连接数等服务质量指标的要求。未来的服务场景会更加灵活多变, 涉及多种服务指标的组合甚至动态组合。另外, 不同网络服务或是同一网络服务的不同阶段, 也会对网络功能提出不同的需求。因此, 知识定义多模态网络考虑使用一组网络指标向量和一组网络功能向量, 来对服务场景进行细粒度的描述和刻画。

2) 网络设备的多态化。未来网络的设备将呈现“端-边-云”融合的态势, 不可避免地将出现具有不同网络功能、提供不同网络资源、具有不同配置语言的网络设备共同存在并相互协作、共同提供网络服务的状态。因此, 知识定义多模态网络考虑使用一组网络功能向量、一组网络资源向量和一组配置语言向量来对网络设备进行抽象化的定义。

3) 网络知识的多态化。在知识定义的多模态网络中, 知识空间还涵盖了多个不同层次、不同侧面的知识内容, 并且对这些知识内容之间进行了有机的组合、推理和利用。比如, 描述性的知识可大致

分为服务需求、设备控制和网络组织 3 个部分, 分别对应并归纳网络运行过程中一个侧面的基础信息。在这些描述性的信息之间, 可以通过发掘连接关系、依赖关系、约束关系等方式, 构建出网络知识空间图谱, 然后利用这一图谱在不同的知识组成部分之间进行联动。

网络的知识范畴是广泛而海量的, 不同网络位置之间的知识传递和同步是一个关键环节。知识定义多模态网络可以更进一步地以分布式的形式加以扩展, 如图 4 所示, 在不同网络位置之间通过传递关键的知识特征向量来对整体网络的知识空间进行高效同步。在知识定义多模态网络中, 每个网络智能管理域针对关键知识特征的生成, 采用了图感知神经网络的方法对域内已知的网络知识空间图谱进行处理, 使用注意力机制提取知识空间图谱中对解决网络优化问题最有价值的键信息, 输出关键特征向量以便在域间进行传递。同时, 每个网络智能管理域在接收到其他域传来的关键知识特征之后, 会将该关键知识特征向量

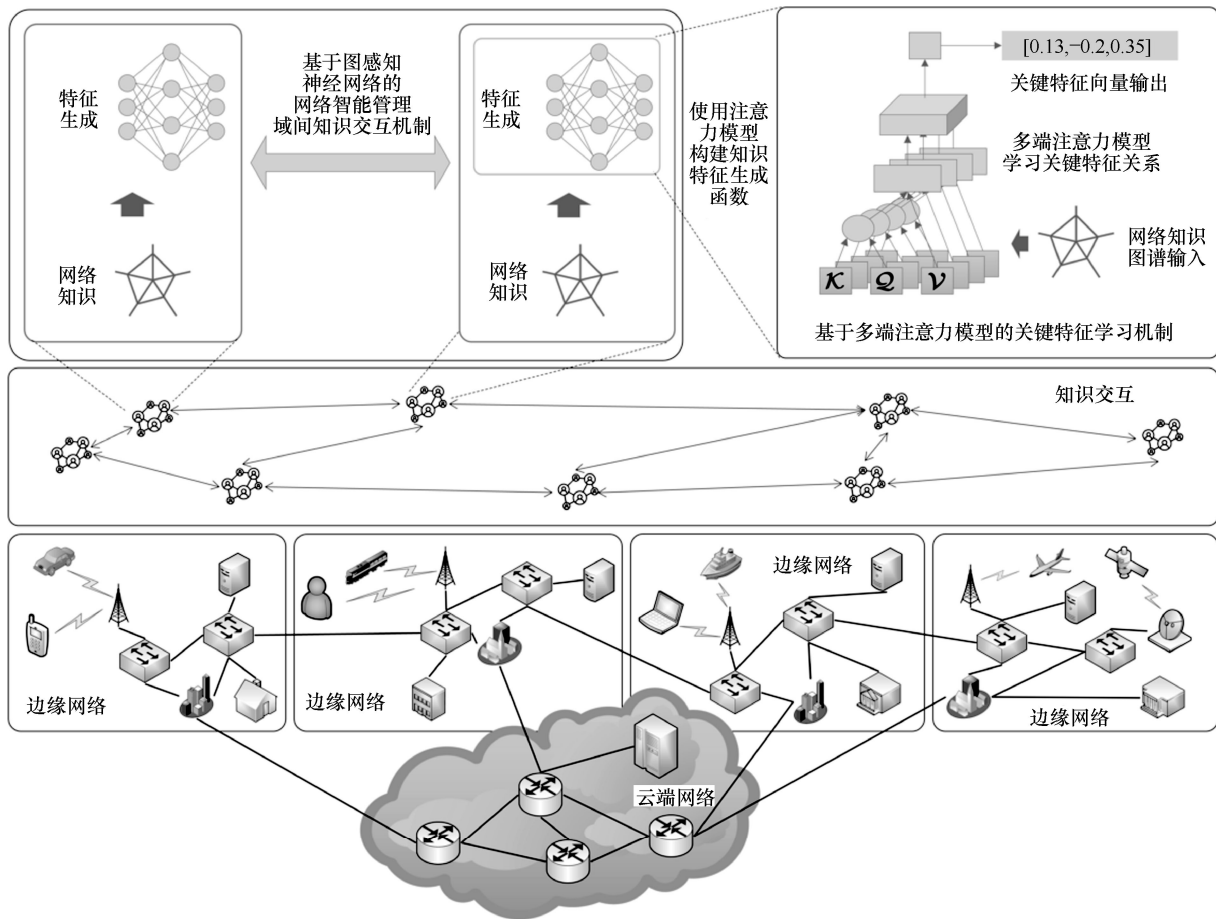


图 4 知识定义多模态网络的域间分布式关键知识交互

更新到本地的已知知识库当中。

在知识定义多模态网络中，差异化的网络知识将能够提升网络的优化和管理能力，最终网络整体性能的提升则依赖于局域网络制式之间的良好交互。当然，这同时也意味着一系列的挑战，在实际应用研究中应加以注意。首先，在局域网络知识的学习、抽取和总结过程当中，依然需要依赖各种人工智能技术。人工智能技术的应用，一方面需要寻找与应用场景、局部网络结构适用的建模方式和学习算法，本文相信强化学习是其中一个重要的关键技术分支，同时需要关注在线/离线式学习的合理组合，以提升对已有案例数据的利用效率，减小学习过程中的事件成本。另一方面，人工智能技术的应用不可避免地对网络提出了更高的算力需求，这需要与边缘计算、边缘网络、通信感知计算一体化等研究方向一同努力，使智能可以真正地根植在网络之中，形成网络“自智”。其次，对于不同局域的网络知识交互，也需要精细化的设计。一方面，消息传递图神经网络将神经网络和通信过程结合到一起，是网络知识交互过程的良好载体；另一方面，为了更高效地利用网络知识，并且根据这些知识对网络优化管理进行推理决策，还需要结合实际网络问题，研究知识蒸馏方法和知识图谱的构建方法。最后，从数学化的网络优化方案的生成，到最终针对网络设备和业务的配置与部署，也还存在距离。在这个问题上，相关关键技术瓶颈与意图网络等研究方向存在交叉，可以相互借鉴、推进。

3 知识定义多模态网络建模

3.1 模型定义

本节首先定义模型的一系列参数，如表1所示。具体地，定义服务质量指标的数量为 n^M ，则有服务质量指标向量 $\mathbf{m} \in \mathbb{R}^{n^M}$ 为

$$\mathbf{m} = (m_1, m_2, \dots, m_{n^M}) \quad (1)$$

以及服务质量评价函数 $p(\mathbf{m}) : \mathbb{R}^{n^M} \rightarrow \mathbb{R}$ 。定义服务功能的数量为 n^F ，则有服务功能向量 $\mathbf{f} \in \mathbb{R}^{n^F}$ 为

$$\mathbf{f} = (f_1, f_2, \dots, f_{n^F}) \quad (2)$$

定义网络资源的数量为 n^R ，则有网络资源向量 $\mathbf{r} \in \mathbb{R}^{n^R}$ 为

$$\mathbf{r} = (r_1, r_2, \dots, r_{n^R}) \quad (3)$$

表1 系统参数

参数	含义
n^M	服务质量指标数量
n^F	服务功能数量
n^R	网络资源数量
n^L	网络设备配置语言数量
n^D	网络需求上限
\mathbf{m}	服务质量指标向量
\mathbf{f}	服务功能向量
\mathbf{r}	网络资源向量
\mathbf{l}	网络设备配置语言向量
$G(V, E)$	网络智能管理域间关系图
$H^v(U, O)$	网络智能管理域内关系图
Ψ^v	域间知识
Θ^v	域内知识
h_i^c	系统隐含状态
k_i^c	知识特征向量
\mathcal{Q}	查询向量
\mathcal{K}	键向量
\mathcal{V}	值向量
$M_i(\cdot)$	知识特征生成函数
$R(\cdot)$	系统性能读出函数

定义网络配置语言数量为 n^L ，则有二进制编码的网络配置语言向量 $\mathbf{l} \in \{0, 1\}^{n^L}$ ，其元素的定义为

$$l_i = \begin{cases} 0, & \text{设备不支持该配置语言} \\ 1, & \text{设备支持该配置语言} \end{cases} \quad (4)$$

定义网络为一个有顶点集 V 和边集 E 组成的图 $G(V, E)$ ，其中每个顶点是一个网络智能管理域，每条边连接着2个连通的智能管理域。同时，每个顶点 v 可以由网络资源向量 \mathbf{r}^v 、网络配置语言向量 \mathbf{l}^v 和边 e 上的域间通信容量限制 c^e 来描述。设图 G 中的顶点数量为 n^G ，域间链路数量为 n^E ，则边集 E 可以表示为

$$E = \{e_1, e_2, \dots, e_{n^E}\} \quad (5)$$

在同一个网络智能管理域 v 内部，同样可以构建一个图 $H^v(U, O)$ ，其中顶点集 U 由具体的网络设备 u 组成。类似地，可以写出设备 u 对应的资源向量 \mathbf{r}^u 。

对于任意一个服务需求 d ，可以使用如下方式定义

$$d = \{v_s, v_t, z, m, f, p\} \quad (6)$$

其中, v_s 是源管理域, v_t 是目标管理域, z 是服务流量需求, m 是服务质量指标, f 是服务功能需求, p 是服务评价函数。令 d_{n^D} 为系统能处理的服务需求数量上限, 则所有服务需求的集合可以表示为

$$D = \{d_1, d_2, \dots, d_{n^D}\} \quad (7)$$

令 x_i 为路由方案在边 e_i 上分配的流量, 则管理域间路由与资源优化的解可以表示为

$$X = \{x_1, x_2, \dots, x_{n^E}\} \quad (8)$$

X 所涉及顶点的指示函数为

$$I(X, v) = \begin{cases} 0, & \neg \exists x_i > 0, \forall e_i, v \in e_i \\ 1, & \exists x_i > 0, \forall e_i, v \in e_i \end{cases} \quad (9)$$

记上标 i 表示需求 d_i 对应的元素, 则优化问题可以定义为

$$\begin{aligned} \max_{x^i, \forall 0 \leq i < n^D} & \sum_i^{n^D} p^i(m^i, z^i, X^i) \\ \text{s.t.} & \sum_i^{n^D} x_j^i \leq c^{e_j}, \forall e_j \in E \\ & \sum_i^{n^D} f_j^i \leq r_j^i, \forall v_j \in V, I(X^i, v_j) = 1 \end{aligned} \quad (10)$$

3.2 知识空间

图 H^v 中所有的资源向量可以组成域内知识 Θ^v 。与此同时, 智能管理域 v 还会维护一个域间知识 Ψ^v

$$\Psi^v = \{k_1, k_2, \dots, k_{n^G}\}^T \quad (11)$$

其中, k_v 是管理域 v 向外部交换的关键知识特征向量, 其基于域内知识 Θ^v 和域间知识 Ψ^v 并使用神经

网络生成。令系统的隐含状态 h'_v 为

$$h'_v = \{\Theta^v, \Psi^v\} \quad (12)$$

知识特征向量由函数 $M_t(\cdot)$ 生成, 即

$$k_v^{t+1} = M_t(h'_v) \quad (13)$$

由于知识特征向量 k_v 沿管理域之间的拓扑关系迭代传递, 其生成机制实质上是一种消息传递图神经网络 (MPGNN, message passing graph neural network)。关键知识特征生成与域间知识传递机制如图 5 所示。

生成函数 $M_t(\cdot)$ 使用多端注意力机制进行实现。对于每一个注意力端, 使用如下的注意力机制

$$\text{Att}(\mathcal{Q}, \mathcal{K}, \mathcal{V}) = \text{softmax}\left(\frac{\mathcal{Q}\mathcal{K}^T}{\sqrt{\text{dim}}}\right)\mathcal{V} \quad (14)$$

其中, \mathcal{Q} 、 \mathcal{K} 和 \mathcal{V} 分别是查询向量、键向量和值向量, dim 是输入特征维度。多端注意力模型将最终的输出分摊给多个注意力模型进行处理, 假设一共有 n^{head} 个注意力端, 令 $W_i^{\mathcal{K}}$ 、 $W_i^{\mathcal{Q}}$ 和 $W_i^{\mathcal{V}}$ 分别对应端 i 中键向量、查询向量和值向量各自的处理矩阵, $W^{\mathcal{O}}$ 是输出向量的处理矩阵。那么多端注意力机制可以表示为

$$\text{MHAtt}(\cdot) = \text{Concat}(\mathcal{H}_1, \dots, \mathcal{H}_{n^{\text{head}}})W^{\mathcal{O}} \quad (15)$$

$$\mathcal{H}_i(\cdot) = \text{Att}(\mathcal{Q}W^{\mathcal{Q}}, \mathcal{K}W^{\mathcal{K}}, \mathcal{V}W^{\mathcal{V}}) \quad (16)$$

更进一步地, 使用自注意力机制^[48]进行实现, 即查询向量、键向量和值向量均来自同一输入。这里使用系统的隐含状态 h'_v 作为输入。

$$\mathcal{H}_i(h'_v) = \text{Att}(h'_vW^{\mathcal{Q}}, h'_vW^{\mathcal{K}}, h'_vW^{\mathcal{V}}) \quad (17)$$

3.3 优化与部署

在每个智能管理域内, 可以使用读出函数 $R(\cdot)$

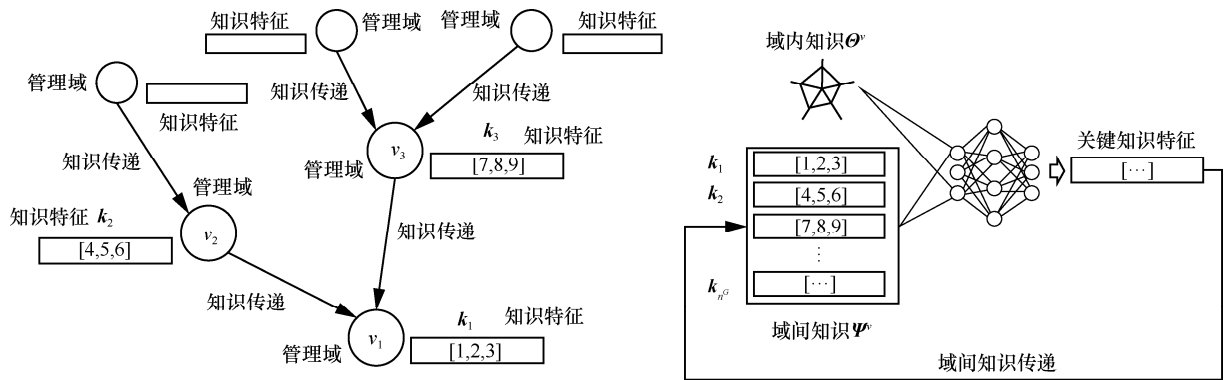


图 5 关键知识特征生成与域间知识传递机制

来拟合 $\mathbb{E}\left[\sum_i^{n^D} p^i(m^i, z^i, X^i)\right]$, 那么在 t 时刻, 系统的解决方案求解可以表示为

$$X = \arg \max_X R(k_v^t, X) \quad (18)$$

由于每个智能管理域所支持的配置语言 I 的不同, 优化方案 X 可以被实现为不同的部署方案, 记该部署方案为 S , 则有

$$S = T(G, I, X) \quad (19)$$

其中, T 是部署生成函数, 可以考虑使用意图网络 (IBN, intent-based network) [49-52]、语义图对齐 [53] 等技术进行实现。

令网络的运行环境函数为 ε , 则在部署方案 S 下的实际运行结果可以表示为

$$Y = \varepsilon(G, S) \quad (20)$$

因此, 可以使用深度强化学习方法来寻找一个值函数作为读出函数 $R(\bullet)$ 来拟合长期收益

$$\mathbb{E}\left[\sum_i^{n^D} p^i\left(m^i, z^i, \varepsilon\left(G, T\left(G, I, \arg \max_{X^i} R^i(k_v^t, X^i)\right)\right)\right)\right] \quad (21)$$

其中, 面向不同需求的读出函数可以考虑使用多任务强化学习方法进行联合实现 [54-56]。

进一步地, 式(10)中定义的优化问题可以变换为

$$\begin{aligned} & \max_{X^i, \forall 0 \leq i < n^D} \mathbb{E}\left[\sum_i^{n^D} p^i(m^i, z^i, Y^i)\right] \\ & \text{s.t.} \quad \sum_i^{n^D} x_j^i \leq c^{e_j}, \forall e_j \in E \\ & \quad \sum_i^{n^D} f_j^i \leq r_j^i, \forall v_j \in V, I(X^i, v_j) = 1 \\ & \quad X^i = \arg \max_{X^i} R^i(k_v^t, X^i) \\ & \quad S^i = T(G, I, X^i) \\ & \quad Y^i = \varepsilon(G, S^i) \end{aligned} \quad (22)$$

4 应用案例

考虑一个多智能管理域的异构网络, 如图6所示。其中, 每个管理域包含2种节点, 传统节点只能根据目的地址确定数据流转路径, 而智能节点则可以根据更精细的匹配规则来确定数据流的转

发路径。在不同管理域内, 由于设备的区别, 其智能节点支持的匹配级别也不一致。每个节点都有流向其他节点的流量需求, 每个管理域预先定义了100个流量模式, 并且流量模式之间的切换按照预先定义的概率进行选择。

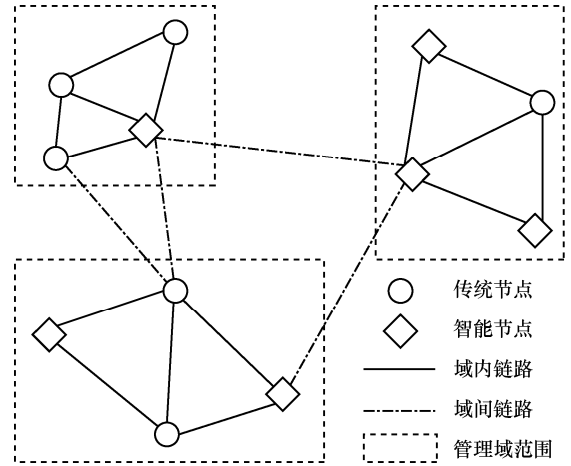


图6 多智能管理域异构网络示意

下面, 以求解异构场景下的流量工程问题作为知识定义多模态网络的应用案例。如图7所示, 每个智能管理域构建一个任务专用的神经网络, 其输出表示给定任务需求时选择对应路径的概率。每20个时间片为一个时间阶段, 当前时间阶段的最后一个时间片获得的立即回报值为前一时间阶段最大链路利用率减去当前时间阶段最大链路利用率的差值。本文使用值迭代的方法修正路径的长期回报 Q , 并根据梯度修正路径选择神经网络的参数, 生成新的路径选择方案模型。

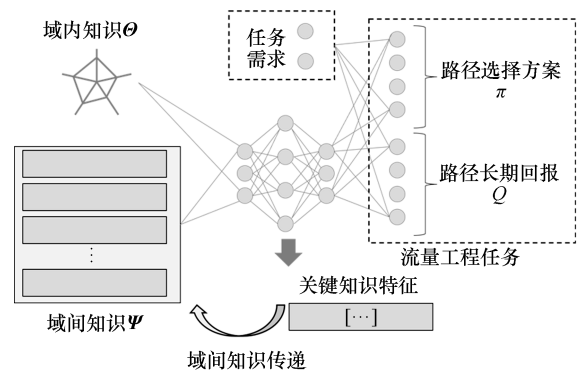


图7 知识定义多模态网络在流量工程任务中的应用

仿真实验与一系列基准算法进行了对比, 分别是原始的最短路径选路 (SPF, shortest first path) 方法、基于线性规划的流量工程 (TE, traffic engineering) 方

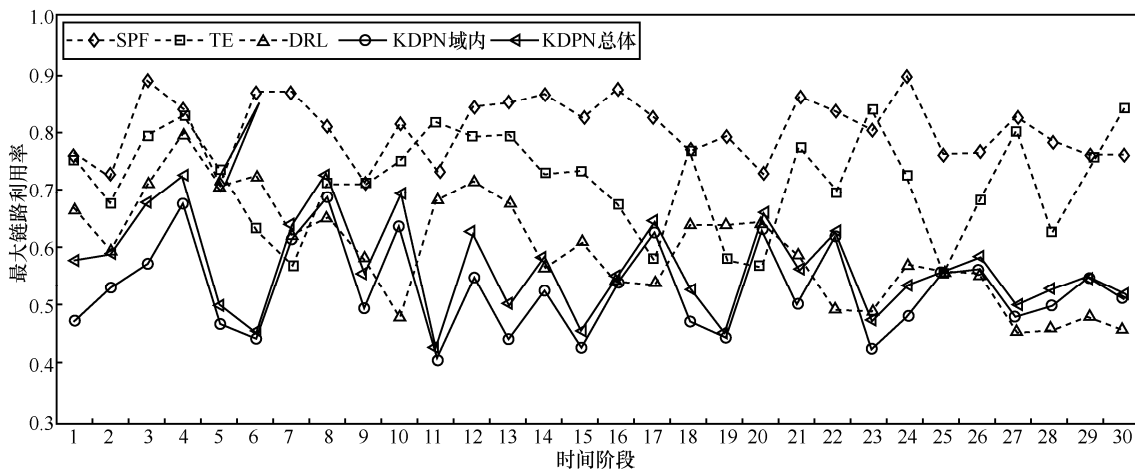


图 8 最大链路利用率的变化

法、基于全局深度强化学习 (DRL, deep reinforcement learning) 模型的方法。本文的知识定义多模态框架下的 KDPN 方法分为 2 个数据结果, 一个是只考虑域内流量的结果, 另一个是同时考虑域内和域间流量的总体结果。实验过程中最大链路利用率的变化如图 8 所示, KDPN 的域内流量收敛速度更有优势; 同时从表 2 对统计指标的对比中可以看到, KDPN 与基准算法相比取得了更好的效果。

表 2 仿真实验结果统计指标

算法	平均值	标准差
SPF	0.806 0	0.055 2
TE	0.716 2	0.086 0
DRL	0.599 2	0.090 8
KDPN 域内	0.530 6	0.076 6
KDPN 总体	0.568 3	0.081 3

5 结束语

本文为了解决面向万物互联场景下移动通信网络用户复杂多变的需求的业务网络能力提供问题, 针对管理、控制、数据三平面解耦的未来移动通信网络系统, 将“网络知识”作为贯穿多个平面的“基因”主干, 提出了一种知识定义多模态网络按需服务体系架构, 支持使用同一网络体系架构上, 面向不同服务场景、跨越多级服务层次、融合多种服务指标, 为不特定的最终用户提供移动通信业务网络服务, 使用户的需求可以得到及时、有效的满足和保障。与此同时, 管理、控制、数据三平面解耦是通信网络发展的一个重要趋势, 本文所提

架构对于更广义、普适的通信网络中差异化的网络优化与管理也有借鉴价值。

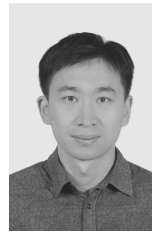
参考文献:

- [1] CHEN X, JIAO L, LI W Z, et al. Efficient multi-user computation offloading for mobile-edge cloud computing[J]. IEEE/ACM Transactions on Networking, 2016, 24(5): 2795-2808.
- [2] SARDELLITTI S, SCUTARI G, BARBAROSSA S. Joint optimization of radio and computational resources for multicell mobile-edge computing[J]. IEEE Transactions on Signal and Information Processing Over Networks, 2015, 1(2): 89-103.
- [3] 张建敏, 谢伟良, 杨峰义, 等. 移动边缘计算技术及其本地分流方案[J]. 电信科学, 2016, 32(7): 132-139.
ZHANG J M, XIE W L, YANG F Y, et al. Mobile edge computing and application in traffic offloading[J]. Telecommunications Science, 2016, 32(7): 132-139.
- [4] LIANG B, WONG V, SCHOBBER, R, et al. Mobile edge computing[J]. Key Technologies for 5G Wireless Systems, 2017, 16(3): 1397-1411.
- [5] TALEB T, SAMDANIS K, MADA B, et al. On multi-access edge computing: a survey of the emerging 5G network edge cloud architecture and orchestration[J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1657-1681.
- [6] MAO Y Y, ZHANG J, SONG S H, et al. Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems[J]. IEEE Transactions on Wireless Communications, 2017, 16(9): 5994-6009.
- [7] AI Y, PENG M G, ZHANG K C. Edge computing technologies for Internet of things: a primer[J]. Digital Communications and Networks, 2018, 4(2): 77-86.
- [8] 施巍松, 孙辉, 曹杰, 等. 边缘计算: 万物互联时代新型计算模型[J]. 计算机研究与发展, 2017, 54(5): 907-924.
SHI W S, SUN H, CAO J, et al. Edge computing—an emerging computing model for the Internet of everything era[J]. Journal of Computer Research and Development, 2017, 54(5): 907-924.
- [9] CHEN M, HAO Y X. Task offloading for mobile edge computing in

- software defined ultra-dense network[J]. *IEEE Journal on Selected Areas in Communications*, 2018, 36(3): 587-597.
- [10] 李子姝, 谢人超, 孙礼, 等. 移动边缘计算综述[J]. *电信科学*, 2018, 34(1): 87-101.
- LI Z S, XIE R C, SUN L, et al. A survey of mobile edge computing[J]. *Telecommunications Science*, 2018, 34(1): 87-101.
- [11] WANG X F, HAN Y W, WANG C Y, et al. In-edge AI: intelligentizing mobile edge computing, caching and communication by federated learning[J]. *IEEE Network*, 2019, 33(5): 156-165.
- [12] POULARAKIS K, LLORCA J, TULINO A M, et al. Joint service placement and request routing in multi-cell mobile edge computing networks[C]//*Proceedings of IEEE Conference on Computer Communications*. Piscataway: IEEE Press, 2019: 10-18.
- [13] 施巍松, 张星洲, 王一帆, 等. 边缘计算: 现状与展望[J]. *计算机研究与发展*, 2019, 56(1): 69-89.
- SHI W S, ZHANG X Z, WANG Y F, et al. Edge computing: state-of-the-art and future directions[J]. *Journal of Computer Research and Development*, 2019, 56(1): 69-89.
- [14] KAMEL M, HAMOUDA W, YOUSSEF A. Ultra-dense networks: a survey[J]. *IEEE Communications Surveys & Tutorials*, 2016, 18(4): 2522-2545.
- [15] RUI L L, YANG Y T, GAO Z P, et al. Computation offloading in a mobile edge communication network: a joint transmission delay and energy consumption dynamic awareness mechanism[J]. *IEEE Internet of Things Journal*, 2019, 6(6): 10546-10559.
- [16] 张海波, 李虎, 陈善学, 等. 超密集网络中基于移动边缘计算的任务卸载和资源优化[J]. *电子与信息学报*, 2019, 41(5): 1194-1201.
- ZHANG H B, LI H, CHEN S X, et al. Computing offloading and resource optimization in ultra-dense networks with mobile edge computation[J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1194-1201.
- [17] CLAFFY K, CLARK D. Platform models for sustainable Internet regulation[J]. *Journal of Information Policy*, 2014, 4(1): 463-488.
- [18] LEHR W, CLARK D D, BAUER S, et al. Regulation when platforms are layered[J]. *Social Science Electronic Publishing*, 2019: doi.org/10.2139/ssrn.3427499.
- [19] GIOTSAS V, NOMIKOS G, KOTRONIS V, et al. O peer, where art thou? uncovering remote peering interconnections at IXPs[J]. *IEEE/ACM Transactions on Networking*, 2021, 29(1): 1-16.
- [20] JIANG W, STRUFE M, SCHOTTEN H D. Experimental results for artificial intelligence-based self-organized 5G networks[C]//*Proceedings of 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*. Piscataway: IEEE Press, 2017: 1-6.
- [21] KIBRIA M G, NGUYEN K, VILLARDI G P, et al. Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks[J]. *IEEE Access*, 2018, 6: 32328-32338.
- [22] PARK J, SAMARAKOON S, BENNIS M, et al. Wireless network intelligence at the edge[J]. *Proceedings of the IEEE*, 2019, 107(11): 2204-2239.
- [23] XU S J, QIAN Y, HU R Q. Data-driven network intelligence for anomaly detection[J]. *IEEE Network*, 2019, 33(3): 88-95.
- [24] XU S J, QIAN Y, HU R Q. Data-driven edge intelligence for robust network anomaly detection[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 7(3): 1481-1492.
- [25] GUTIERREZ-ESTEVEZ D M, GRAMAGLIA M, DOMENICO A D, et al. Artificial intelligence for elastic management and orchestration of 5G networks[J]. *IEEE Wireless Communications*, 2019, 26(5): 134-141.
- [26] CHEN D W, LIU Y C, KIM B, et al. Edge computing resources reservation in vehicular networks: a meta-learning approach[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(5): 5634-5646.
- [27] QIAO C M, MEI Y S, MYUNGSIK Y, et al. Polymorphic control for cost-effective design of optical networks[J]. *European Transactions on Telecommunications*, 2000, 11(1): 17-26.
- [28] CHEN L, MNAOUER A B, FOH C H. An optimized polymorphic hybrid multicast routing protocol (OPHMR) for ad hoc networks[C]//*Proceedings of 2006 IEEE International Conference on Communications*. Piscataway: IEEE Press, 2006: 3572-3577.
- [29] 胡宇翔, 董芳, 王鹏, 等. 面向多样化服务定制的多态路由机制研究[J]. *通信学报*, 2015, 36(7): 48-59.
- HU Y X, DONG F, WANG P, et al. Research on polymorphic routing mechanism for customized diversified services[J]. *Journal on Communications*, 2015, 36(7): 48-59.
- [30] 胡宇翔, 伊鹏, 孙鹏浩, 等. 全维可定义的多模态智慧网络体系研究[J]. *通信学报*, 2019, 40(8): 1-12.
- HU Y X, YI P, SUN P H, et al. Research on the full-dimensional defined polymorphic smart network[J]. *Journal on Communications*, 2019, 40(8): 1-12.
- [31] 邬江兴, 胡宇翔. 网络技术体系与支撑环境分离的发展范式[J]. *信息通信技术与政策*, 2021, 47(8): 1-11.
- WU J X, HU Y X. The development paradigm of separation between network technical system and supporting environment[J]. *Information and Communications Technology and Policy*, 2021, 47(8): 1-11.
- [32] HU Y X, LI D, SUN P H, et al. Polymorphic smart network: an open, flexible and universal architecture for future heterogeneous networks[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 7(4): 2515-2525.
- [33] MARDER E. Invertebrate neurobiology: polymorphic neural networks[J]. *Current Biology*, 1994, 4(8): 752-754.
- [34] TEAMA S. DNA polymorphisms: DNA-based molecular markers and their application in medicine[J]. *Genetic Diversity and Disease Susceptibility*, 2019: doi.org/10.1007/978-1-4615-3690-1.
- [35] CLARK D D, PARTRIDGE C, RAMMING J C, et al. A knowledge plane for the Internet[C]//*Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*. [S.l.:s.n.], 2003: 3-10.
- [36] MESTRES A, RODRIGUEZ-NATAL A, CARNER J, et al. Knowledge-defined networking[J]. *ACM SIGCOMM Computer Communication Review*, 2017, 47(3): 2-10.
- [37] 朱近康. 知识+数据驱动学习: 未来网络智能的基础[J]. *中兴通讯技术*, 2020, 26(4): 46-49.
- ZHU J K. Knowledge-and-data driven learning: foundation of future network intelligence[J]. *ZTE Technology Journal*, 2020, 26(4): 46-49.
- [38] HYUN J, TU N V, HONG J W K. Towards knowledge-defined networking using in-band network telemetry[C]//*Proceedings of 2018 IEEE/FIP Network Operations and Management Symposium*. Piscataway: IEEE Press, 2018: 1-7.

- [39] PHAM T A Q, HADJADJ-AOUL Y, OUTTAGARTS A. Deep reinforcement learning based qos-aware routing in knowledge-defined networking[C]//Proceedings of the International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness. Berlin: Springer, 2018: 14-26.
- [40] XIE J F, YU F R, HUANG T, et al. A survey of machine learning techniques applied to software defined networking (SDN): research issues and challenges[J]. IEEE Communications Surveys & Tutorials, 2019, 21(1): 393-430.
- [41] GARRAHAN J J, RUSSO P A, KITAMI K, et al. Intelligent network overview[J]. IEEE Communications Magazine, 1993, 31(3): 30-36.
- [42] MANOJ B S, RAO R R, ZORZI M. CogNet: a cognitive complete knowledge network system[J]. IEEE Wireless Communications, 2008, 15(6): 81-88.
- [43] YANG S Y, CHANG Y Y. An active and intelligent network management system with ontology-based and multi-agent techniques[J]. Expert Systems with Applications, 2011, 38(8): 10320-10342.
- [44] JIANG W, STRUFE M, SCHOTTEN H D. Intelligent network management for 5G systems: the SELFNET approach[C]//Proceedings of 2017 European Conference on Networks and Communications (EuCNC). Piscataway: IEEE Press, 2017: 1-5.
- [45] QI Q, WANG J Y, MA Z Y, et al. Knowledge-driven service offloading decision for vehicular edge computing: a deep reinforcement learning approach[J]. IEEE Transactions on Vehicular Technology, 2019, 68(5): 4192-4203.
- [46] DONG T J, QI Q, WANG J Y, et al. Generative adversarial network-based transfer reinforcement learning for routing with prior knowledge[J]. IEEE Transactions on Network and Service Management, 2021, 18(2): 1673-1689.
- [47] WEI Y F, YU F R, SONG M, et al. Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor-critic deep reinforcement learning[J]. IEEE Internet of Things Journal, 2019, 6(2): 2061-2073.
- [48] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the Advances in Neural Information Processing Systems. [S.l.:s.n.], 2017: 5998-6008.
- [49] SZIGETI T, ZACKS D, FALKNER M, et al. Cisco digital network architecture: intent-based networking for the enterprise[M]. S.l.: Cisco Press, 2018.
- [50] RIFTADI M, KUIPERS F. P4I/O: intent-based networking with P4[C]//Proceedings of 2019 IEEE Conference on Network Softwarization (NetSoft). Piscataway: IEEE Press, 2019: 438-443.
- [51] 李福亮, 范广宇, 王兴伟, 等. 基于意图的网络研究综述[J]. 软件学报, 2020, 31(8): 2574-2587.
- LI F L, FANG Y, WANG X W, et al. State-of-the-art survey of intent-based networking[J]. Journal of Software, 2020, 31(8): 2574-2587.
- [52] 王敬宇, 周铖, 张蕾, 等. 知识定义的意图网络自治[J]. 电信科学, 2021, 37(9): 1-13.
- WANG J Y, ZHOU C, ZHANG L, et al. Knowledge-defined intent-based network autonomy[J]. Telecommunications Science, 2021, 37(9): 1-13.
- [53] TANG W, WANG J Y, QI Q, et al. Deep graph alignment network[J]. Neurocomputing, 2021, 465: 289-300.
- [54] LIU C H, DAI Z P, YANG H M, et al. Multi-task-oriented vehicular crowdsensing: a deep learning approach[C]//Proceedings of IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2020: 1123-1132.
- [55] QIAN L P, WU Y, JIANG F L, et al. NOMA assisted multi-task multi-access mobile edge computing via deep reinforcement learning for industrial Internet of things[J]. IEEE Transactions on Industrial Informatics, 2021, 17(8): 5688-5698.
- [56] QI Q, ZHANG L X, WANG J Y, et al. Scalable parallel task scheduling for autonomous driving using multi-task deep reinforcement learning[J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 13861-13874.

[作者简介]



王敬宇 (1978-), 男, 吉林长春人, 博士, 北京邮电大学教授、博士生导师, 主要研究方向为智能网络、机器学习、边缘计算等。



庄子睿 (1993-), 男, 北京人, 博士, 北京邮电大学在站博士后, 主要研究方向为网络智能路由、资源优化、深度强化学习、图神经网络等。