



专题：智能通信技术

跨域场景下的联邦路由机制设计

丛培壮¹, 张宇超¹, 田野¹, 王文东¹, 李丹²

(1. 北京邮电大学, 北京 100876; 2. 清华大学, 北京 100084)

摘 要: 随着多网融合的发展, 独立自治域网络之间的高效互联互通已成为一个关键理论技术问题。由于域内信息的私密性, 各自域之间不会共享相关的数据, 传统的域间路由协议无法提供高效的跨域路由机制。为克服现有协议存在的不足, 提出了一种基于机器学习的联邦路由机制, 该机制通过各自域的神经网络模型和参数隐式地共享域内信息, 不仅可打破现存的数据孤岛问题, 还可极大地减少域间信息共享所需要传输的数据量, 进而降低全网路由信息收敛时延, 基于该联邦路由机制, 边界路由器也能够根据全网状态信息制定全局的优化路由转发策略。

关键词: 跨域网络; 联邦路由; 网络融合

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2020286

Design of federated routing mechanism in cross-domain scenario

CONG Peizhuang¹, ZHANG Yuchao¹, TIAN Ye¹, WANG Wendong¹, LI Dan²

1. Beijing University of Posts and Telecommunications, Beijing 100876, China

2. Tsinghua University, Beijing 100084, China

Abstract: With the development of multi-network integration, how to ensure efficient interconnections among multiple independent network domains is becoming a key problem. Traditional interdomain routing protocol fails due to the limitation of domain information privacy, where each autonomous domain doesn't share any specific intra-domain information. A machine learning-based federated routing mechanism was proposed to overcome the existing shortcomings. Each autonomous domain shares intra-domain information implicitly via neural network models and parameters. It not only breaks data islands problems but also greatly reduces the amount of transmitted data shared between domains, then decreases convergence delay of entire network information. Based on the federated routing mechanism, border routers can formulate global optimal routing strategies according to the status of entire network.

Key words: cross-domain network, federated routing, network integration

收稿日期: 2020-09-02; 修回日期: 2020-10-10

通信作者: 张宇超, yczhang@bupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61802024); 中央高校基本科研业务费专项资金资助项目 (No.2482020RC36); 国家重点研发计划项目 (No.2019YFB1802603)

Foundation Items: The National Natural Science Foundation of China (NSFC) Youth Science Foundation (No.61802024), The Fundamental Research Funds for The Central Universities under Grant (No.2482020RC36), The National Key Research and Development Program of China (No.2019YFB1802603)



1 引言

随着网络技术的飞速发展,越来越多的设备得以互联,通常处于同一个管理机构控制之下的路由器和网络群组会组成一个 AS (autonomous system, 自治系统)。其中,自治系统可以是一个网络服务商、一所学校或者一个公司个体,同时也被称为一个路由域 (routing domain)。域内所有的路由器互联并通过运行相同的 IGP (interior gateway protocol, 内部网关协议) 交换路由信息,如 RIP (routing information protocol, 路由信息协议)、OSPF (open shortest path first, 开放式最短路径优先)、IS-IS (intermediate system-to-intermediate system, 中间系统到中间系统)、IGRP (interior gateway routing protocol, 内部网关路由协议)、EIGRP (enhanced interior gateway routing protocol, 增强内部网关路由线路协议) 等。同时,域间路由也是互联网路由体系结构的重要构成部分,自治域的边界路由器通过域间路由协议来交换所需要的路由信息,但是域间路由具有管控难度大、路由因素复杂等特点。BGP (border gateway protocol, 边界网关协议) 是当前应用最为广泛的域间路由协议,其可以提供灵活丰富的路由策略完成流量路径规划和负载均衡,但是随着网络流量的激增以及业务需求的多样化,网络运维和管理也随之增加^[1]。此外,BGP 的实现规则使得其存在着不可避免且互相矛盾的问题。

- 数据孤岛: 各自治系统鉴于数据的隐私性并不会提供完全的域内信息,这使得丧失网络全局视角的路由计算无法实现全局最优的跨域路由策略。
- 数据量庞大: 各自治系统交换的路由信息越详尽,计算的路由策略越能够接近最优,但是这增加了每次路由信息交换所需要传输的数据量,从而增加了路由信息的收敛时延和流失率。

机器学习和人工智能技术已经广泛应用到各个领域之中,并且已经取得了显著的效果,那么是否可以利用相关的技术解决域间路由中存在的问题? 神经网络尤其是深度神经网络对输入的数据经过一系列的运算得到目标值,如果利用神经网络模型对域内的网络状态进行拟合,根据输入的传输需求得到传输路径,那么该网络模型作为黑盒可解决域间路由的数据孤岛问题,而且通过神经网络及其参数很难反推出域内网络的状态,从而保护了域内的敏感数据;网络模型和参数的数据体量较小,同时也可解决域间路由交换数据量庞大的问题。

根据这一主要思路,设计了一种跨域场景下联邦路由机制^[2]。各个自治域向控制器上传本域的路由决策神经网络模型,控制器根据各域的连接状态联邦各网络模型并下发至各域的边界路由器。联邦网络模型具有网络全局视角,边界路由器可以根据联邦网络模型计算跨域的最佳路由路径。该联邦方式不受最佳目标限制条件约束,无论是以最少跳数、传输时延,还是以网络吞吐量作为目标,只要各自自治域上传根据约束条件训练好的模型和参数即可。

2 大规模网络中的路由体系现状

目前存在着大量的大规模数据跨域传输的需求,如气象领域中部署在各地的数据采集点需要跨域汇聚数据,或者华为、腾讯、阿里巴巴等商业公司的各数据中心之间的数据传输等。目前跨域数据传输方案中依然存在着许多不足,因此许多研究机构对跨域场景下的大规模数据传输进行了大量研究。该领域的研究大致可分为两类,一是基于传统网络协议的改进和优化,二是采用基于 SDN (software defined networking, 软件定义网络) 技术的大规模数据传输机制。

2.1 基于传统网络协议的跨域数据传输

BGP 是目前应用最为广泛的域间路由协议,当两个 AS 需要交换路由信息时,每个 AS 都必须

指定一个运行 BGP 的节点即边界网关或边界路由器来代表所属的 AS 与其他 AS 交换路由信息。传统的 BGP 由于缺乏全网拓扑和全局流量观,则只能遵循标准 BGP 路由规则制定路由策略,但随着网络流量的爆发式增长和需求的多样化, BGP 存在的不足日趋明显。在 BGP 的 RR(route reflector, 路由反射器)机制中, RR 只考虑到距离其最近的路由从而产生非最优路由策略。如图 1 所示, 当一个山东用户需要访问韩国运营商的网络, 最佳的策略就是直接访问韩国 PoP (point of presence, 网络接入节点), 而路由反射器能收到两条接入韩国运营商网络的路径, 一是经过北京节点, 二是直接经过韩国 PoP 接入网络。根据路由规则, 路由反射器会反射距离最近的路由, 而 RR 距离北京节点比韩国的 PoP 点更近, 于是把经过北京节点的路径反射给山东节点, 从而该网络传输产生绕路。同时, 传统 BGP 只能匹配目的 IP 地址, 无法满足当前用户或业务传输需要更细力度流量调度的需求。Geng 等^[3]提出了一种基于 BGP 的域间二维路由方案, 该方案通过同时考虑目的地址和源地址实现了灵活、细粒度地流量控制, 并通过二维转发表的方案, 解决二维转发匹配造成的 TCAM (ternary content addressable memory, 三态内容寻址存储器) 空间爆炸的问题。

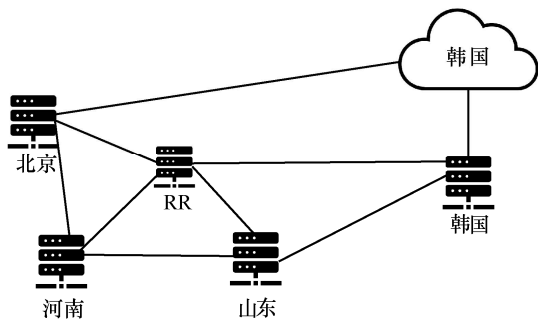


图 1 路由反射机制实例

2.2 基于 SDN 技术的跨域数据传输机制

SDN 技术将数据平面和控制平面解耦, 其提供的可编程接口改变了传统网络的管理方式, 于

是许多研究人员将 SDN 技术应用到跨域大规模网络传输中^[4-5]。Hong 等^[6-7]基于 SDN 技术提出了 SWAN (software-driven wide area network, 软件驱动广域网) 架构, 其工作流程分为以下 3 个步骤: 服务模块将传输任务的如时延、带宽等需求通告给控制器; 控制器根据全局网络拓扑信息和传输需求计算该任务的传输路径; 控制器将传输路径下发至 OpenFlow 交换机。虽然 SWAN 架构较原有传输机制极大提升了传输效率, 但是单一控制器策略一旦产生故障则会严重影响系统的整体效率, 甚至导致系统宕机。Jain 等^[8]提出 B4 架构以实现数据中心的互联, 该架构主要由 Quagga 路由器、SDN 控制器和 OpenFlow 交换机 3 个主要模块组成, 各模块的主要功能如下: Quagga 负责收集域间路由路径信息; SDN 控制器维护全局网络拓扑关系, 并根据 Quagga 收集的路径信息计算出对应的 OpenFlow 流表; OpenFlow 交换机根据流表对数据流进行转发, 但是 B4 中的 Quagga 是根据 BGP 获取路由路径信息的, 所以该架构依然存在着一定的限制。在大规模数据中心场景下, Lv 等^[9]通过优化容器的放置和重分配算法, 在降低通信时延的同时提升了整体服务的吞吐量。Zhang 等^[10]提出一种集中式的架构, 通过中央控制器维护服务器之间数据传输状态的全局视图以充分利用覆盖路径, 极大提升了数据中心关于数据复制任务的传输效率。He 等^[11]在 RR 的基础上提出了基于 SDN 技术的 RR+架构, 在每个 AS 独立设置专属的 SDN 控制器, AS 之间不再通过 BGP 交换信息, 而是通过 SDN 控制器进行数据交换。各 SDN 控制器在收集到全局网络拓扑信息之后再将域间路由由决策下发至本域的所有路由器, 最终 AS 内的路由器根据下发的转发表将需要跨域传输的流量进行转发。Arins 等^[12]提出一种由多个 AS 组成的基于区块链技术的未来网络架构, 其中各 AS 将感知到的链路时延等参数存储至区块链中, 并且各域的 SDN 控制器根据区块链中的可信



数据对数据流进行路由。Xiang 等^[13]提出 SDI (software-defined internet working) 模型, 并在此模型基础上设计了一种基于贝叶斯优化理论的优化算法。该算法通过对端到端路由进行顺序采样, 并通过少量采样找到符合策略的接近最优的端到端路由路径。Zhou 等^[14]为解决基于 SDN 的域间路由机制在信息异步更新时导致的数据分组丢失和服务中断等问题, 基于通过相关域之间的通信和协作, 提出了一种轻量级的、无需扩展交换机南向接口的无缝式域间路由更新方法。Zhang 等^[15]提出了一种 SDNMDR (SDN-based multi-domain routing mechanism, 基于 SDN 的多域路由机制), SDNMDR 通过扩展标准控制器的功能模块, 利用相邻管理域控制器间进行交互与协商, 采用 OpenFlow 协议控制 SDN 交换机, 为跨多管理域的应用提供端到端路径; 并且在各管理域的网络资源允许的情况下, 能够为应用提供 QoS (quality of service, 服务质量) 保证。

2.3 面向跨域网络高效路由的潜在问题

无论是基于传统 BGP 的改进方案, 还是采用 SDN 技术的跨域路由策略, 都依然存在着一些潜在的问题, 并且随着网络用户激增以及网络数据传输需求的不断提升, 潜在问题会愈加突显。其中, 潜在问题可以总结为以下 3 点。

- 互连互通问题: 无论是以网络运营商还是以学校、商业公司等划分的自治域, 考虑到网络的安全性和隐私性, 域内的网络拓扑信息或其他相关数据都应该得以保护, 如果将这些信息扩散至域外, 那么网络安全将受到严重的威胁。每个自治域对信息的保护导致各域之间形成了数据孤岛, 从而导致在跨域路由路径规划时丧失全局视角, 最终无法得到全局最优的路由决策。
- 高时延问题: 根据 BGP, 各域之间需要传输大量路由所需的数据, 这些数据的传输

时延不仅产生时延, 而且也增加网络链路的传输负担。当目的网络不可达或者链路失效时, BGP 会触发重新选路过程, 更新的路由信息则会进一步扩散至其他自治域; 在探索所有可能到达目的网络的路径时可能会导致路由被撤销, 进而增加了全网的路由收敛时间^[16-17]。高时延问题将严重影响路由决策的时效性, 从而影响了跨域网络的数据高效传输。

- 灵活性问题: 默认情况下, BGP 最佳路径都是基于最短自治域的策略。在网络管理中, 诸如链路权重、本地出口优先级以及多出口鉴别器等属性大都基于人为设定^[18], 这也会降低域间路由的灵活性。

3 基于机器学习的联邦路由框架

3.1 联邦路由架构

联邦路由的核心思想是由神经网络以隐式的方法提供域内网络状态或对某路由策略进行评测, 即在获得全局网络状态的同时不对各域内的敏感信息产生安全威胁, 最终计算出最佳的跨域路由路径。基于 SDN 技术的联邦路由架构如图 2 所示。

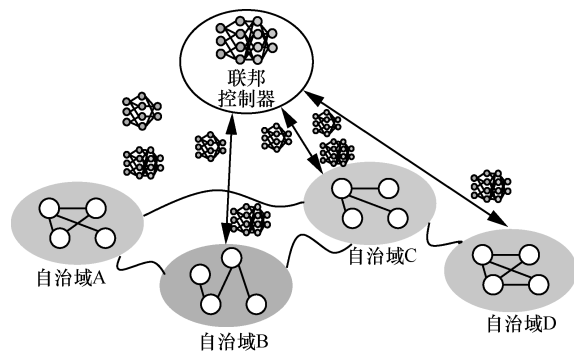


图2 基于 SDN 技术的联邦路由架构

各自治域根据域内的网络拓扑以及链路状态等信息数据训练本地的神经网络模型, 本域神经网络模型根据输入的需求得出域内的最佳路由策略, 其路由评测标准可以根据需求设定, 如转发

跳数、时延或链路利用率等。每个自治域将训练好的模型和参数上传至联邦控制器，联邦控制器结合域间的拓扑状态将上传的域内模型进行联邦，最终再将联邦模型下发至各自治域。当域内网络状态发生变化后需及时更新本域的网络参数，联邦控制器负责更新联邦模型并执行下发操作。为保证对原有网络架构的兼容性，联邦路由机制只需增加域内控制器来负责本域数据采集、模型训练、上传以及接收联邦控制器下发的联邦模型，并将计算得到的跨域路由表下发至域内的边界路由器，尤其是对于不同管理域的网络，为了保证各自治域的最大自主权，各自治域的神经网络模型只需要按照需求提供输入和输出接口。跨地理域网络可以采用多智能体的强化学习进行路由决策^[17]，各自治域中设置智能体（agent），通过联邦各智能体中的策略网络（policy network）和价值网络（value network）计算跨域的全网最佳路由路径。

3.2 联邦路由机制中的关键技术

- 基于神经网络的最短路径策略：神经网络的主要作用是隐式提供网络拓扑及链路状态数据，将域内信息抽象成一个黑盒，对传输需求提供路由策略。强大的神经网络模型可以满足多种需求的传输任务，不再仅限于传统 BGP 基于自治域跳数最短路径计算法原则，可根据需求充分利用域内链路资源和全局网络信息，提供以时延、链路利用路、负载均衡、网络吞吐量以及 QoS 为约束条件的路由策略。基于神经网络的最短路径求解模型如图 3 所示，输入转发需求向量，经过模型运算得出路由路径，结合 SR（segment routing，分段路由）技术完成转发。其中输入的需求向量中应包含目的地址等必要信息，如以网络吞吐量为优化目标则需提供网络流量矩阵等参数。

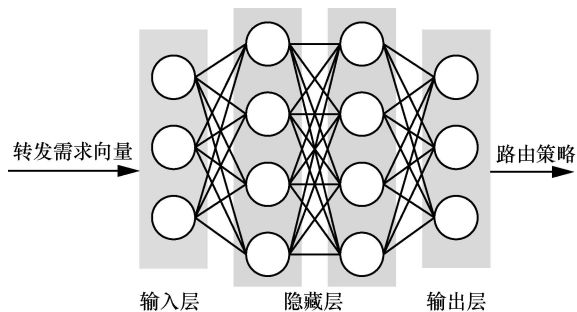
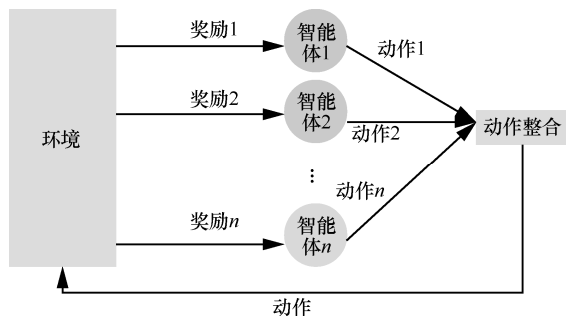


图3 基于神经网络的最短路径策略

- 多智能体强化学习：强化学习非常适用解决决策类问题^[20-21]，在自治域内使用强化学习进行路由决策已经取得了一定的进展^[22-23]，多智能体在其他领域也得到了广泛应用^[24-25]。基本的多智能体强化学习架构如图 4 所示，每个智能体根据环境状态（environment）得出相应的动作（action），整合所有动作并在环境中执行；每个智能体由环境的变化得到对应的奖励（reward），并根据奖励调整下一周期的执行动作。在跨域网络场景下，智能体负责每个自治域，并根据所在域网络状态进行动作决策，将所有动作联邦作为最终的路由决策。环境根据路由策略得到总奖励值并按比例反馈给智能体，以最大化奖励值为目标调整动作决策，模型经过训练最终可得到全局最优的路由路径。



3.3 联邦路由机制优势分析

- 打破数据壁垒：所有自治域的网络拓扑和链路状态等私有信息不必再以显示的方式



进行交换，各自治域的内部数据得以保护，保障了网络的安全性。而且域中负责跨域转发的边界路由器等网络设备具有了全网信息，可以结合传输需求计算出全局最优的跨域路由策略。

- 低时延：传统方式需要交换大量的路由表信息，传输的数据量庞大产生较大的传输时延。联邦路由机制使得各自治域之间只要传输神经网络训练好的模型参数，该部分的数据体量很小，极大地降低了每次需要交换的数据量和传输时延。
- 灵活稳定：功能强大的神经网络可以提供多种约束条件的路由策略，可以是转发跳数、传输时延、链路利用率或者网络的吞吐量等，能够很好地满足广泛的业务传输需求。本地的链路变化在本地即可及时通过模型进行处理，这使得整个联邦路由机制更加稳定。

4 典型业务场景的解决方案

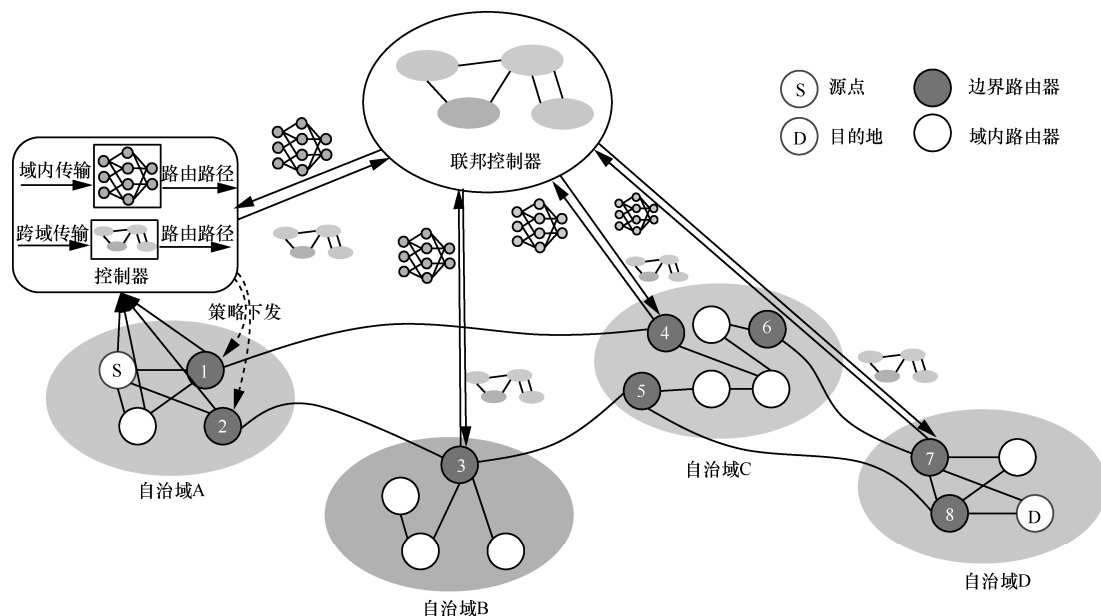
4.1 跨管理域网络中的联邦路由

在跨管理域的网络中，各域需要训练本地的

最短路径求解模型并上传至联邦控制器，联邦控制器结合各域的连通状态对接收的各域内网络模型进行联邦，各自治域的边界路由器根据联邦模型对跨域流量进行转发。一个简单的业务示例如图 5 所示，其中自治域 A 采用额外增加控制器的方式实现联邦路由机制，其余各域选举一个计算能力较强的边缘路由器负责本地模型的训练和上传、路由计算和路由规则下发等操作。假设有数据需要从源点 S 跨域传输至目的地 D，并提出最短转发跳数的传输需求，按传统 BGP 根据最短自治域原则则会从自治域 A 经过自治域 C 到达自治域 D，期间需要经由边界路由器 1、4、6、7，共需要 7 跳到达目的地 D；在联邦路由机制下，则会选择由边界路由器 2、3、5、8，此时虽然经过的自治域增多，但总共却只需 5 跳就能到达目的地。此外，还可以根据链路利用率、传输时延、带宽等属性对域内网络链路设置权值，以灵活满足各种传输需求。

4.2 跨地理域网络中的联邦路由

在跨地理域网络中使用多智能体强化学习模型进行路由决策，采用多智能体的深度确定性策略梯度模型，简要实例如图 6 所示。每个自治域中维护一个智能体，智能体以所属域的网络状态



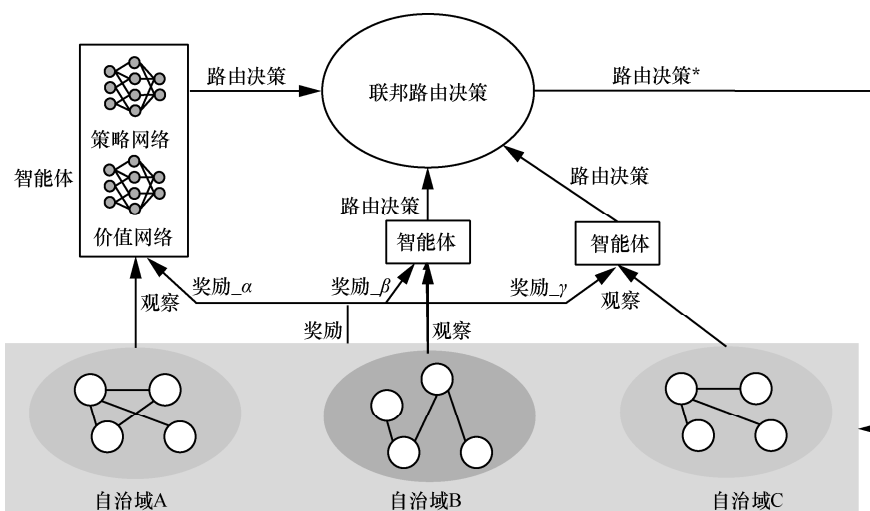


图6 跨地理域网络中的联邦路由机制

作为独立观察环境，并由策略网络、价值网络分别制定、评估决策。联邦控制器将全部路由决策进行联邦并下发至相关的网络控制器进行转发；根据传输结果计算整体奖励和各智能体分属奖励，通过向最大化整体奖励调整各路由决策以获得全局最佳路由路径。

5 结束语

经过对当前大规模跨域网络传输的现状调研，总结出当前跨域路由中存在着难以互联互通、高时延和灵活性不足的问题。本文设计的一种基于机器学习的跨域联邦路由机制能够打破数据壁垒、降低协议数据传输时延和数据量，并提高路由机制的灵活性，从而能够应对更为广泛的传输任务需求。基于机器学习的跨域联邦路由机制是一种全新的路由架构，大量的技术细节和协议部署仍需要进一步的研究和完善。

参考文献:

- [1] TIAN Z, SU S, SHI W, et al. A data-driven method for future Internet route decision modeling[J]. Future Generation Computer Systems, 2019(95): 212-220.
- [2] ZHANG Y C, TIAN Y, WANG W D, et al. Federated routing scheme for large-scale cross domain network[C]//Proceedings of IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscata-

way: IEEE Press, 2020: 1358-1359.

- [3] 耿男, 金飞蔡, 徐明伟. 基于 BGP 的域间二维路由方案[J]. 中兴通讯技术, 2018, 24(6): 64-68.
GENG N, JIN F C, XU M W. A design of inter-domain destination and source routing based on BGP[J]. ZTE Communications, 2018, 24(6): 64-68.
- [4] 赵辉, 丁鸣, 程青松, 等. SDN 与 NFV 技术在云数据中心的规模应用[J]. 电信科学, 2016, 32(1): 144-151.
ZHAO H, DING M, CHENG Q S, et al. Application of SDN and NFV technology in the cloud data center[J]. Telecommunications Science, 2016, 32(1): 144-151.
- [5] 顾戎, 王瑞雪, 李晨, 等. 云数据中心 SDN/NFV 组网方案、测试及问题分析[J]. 电信科学, 2016, 32(1): 126-130.
GU R, WANG R X, LI C, et al. Analysis on network scheme and resolution test of SDN/NFV technology co-deployed in cloud datacenter[J]. Telecommunications Science, 2016, 32(1): 126-130.
- [6] HONG C Y, KANDULA S, MAHAJAN R, et al. Achieving high utilization with software-driven WAN[J]. Computer Communication Review, 2013, 43(4): 5-26.
- [7] MAHAJAN R, HONG C Y, KANDULA S, et al. Achieving high utilization with software-driven WAN(extended version)[R]. 2013.
- [8] JAIN S, KUMAR A, MANDAL S, et al. B4: experience with a globally-deployed software defined WAN[J]. ACM SIGCOMM Computer Communication Review, 2013, 43(4): 3-14.
- [9] LV L, ZHANG Y, LI Y, et al. Communication-aware container placement and reassignment in large-scale internet data centers[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(3): 540-555.
- [10] ZHANG Y, JIANG J, XU K, et al. BDS: a centralized near-optimal overlay network for inter-datacenter data replication[C]//Proceedings of the Thirteenth EuroSys Conference.

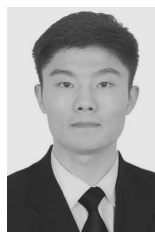


- Porto: [s.n.], 2018: 1-14.
- [11] 何晓明, 刘宁芳, 陈文华. 基于 SDN 的互联网域间路由研究[J]. 通信技术, 2020, 53(5): 1146-1150.
- HE X M, LIU N F, CHEN W H. Research on inter-domain routing based on SDN[J]. Communications Technology, 2020, 53(5): 1146-1150.
- [12] ARINS A. Blockchain based inter-domain latency aware routing proposal in software defined network[C]//Proceedings of 2018 IEEE 6th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE). Piscataway: IEEE Press, 2018: 1-2.
- [13] XIANG Q, ZHANG J, GAO K, et al. Toward optimal software-defined interdomain routing[C]//Proceedings of IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2020: 1529-1538.
- [14] ZHOU H, WU C, CHENG Q, et al. SDN-LIRU: a lossless and seamless method for SDN inter-domain route updates[J]. IEEE/ACM Transactions on Networking, 2017, 25(4): 2473-2483.
- [15] 张俊, 沈苏彬. 一种基于 SDN 的多管理域路由机制[J]. 计算机技术与发展, 2018(8): 86-90.
- ZHANG J, SHEN S B. A SDN-based multi-domain routing mechanism[J]. Computer Technology and Development, 2018(8): 86-90.
- [16] 李琦, 徐明伟, 吴建平. 一种互联网的稳定路由选择策略[J]. 计算机学报, 2012, 35(12): 2668-2675.
- LI Q, XU M W, WU J P. A stable routing selection scheme in the internet[J]. Chinese Journal of Computers, 2012, 35(12): 2668-2675.
- [17] 王洪君, 苑卫鑫, 刘毅, 等. BGP 路由收敛问题及其解决方案研究[J]. 吉林师范大学学报(自然科学版), 2015(2): 125-127.
- WANG H J, YUAN W X, LIU Y, et al. Research of BGP routing convergence problem and solution[J]. Jilin Normal University Journal(Natural Science Edition), 2015(2): 125-127.
- [18] KEHOE M. BGP—the backbone of the Internet[Z]. 2019.
- [19] LOWE R, WU Y, TAMAR A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments[J]. arXiv: 1706.02275, 2017.
- [20] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Massachusetts: MIT Press, 2018.
- [21] NADIGER C, KUMAR A, ABDELHAK S. Federated reinforcement learning for fast personalization[C]//Proceedings of 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE). Piscataway: IEEE Press, 2019: 123-127.
- [22] MAO B, FADLULLAH Z M, TANG F, et al. Routing or computing? The paradigm shift towards intelligent computer net-

work packet transmission based on deep learning[J]. IEEE Transactions on Computers, 2017, 66(11): 1946-1960.

- [23] JIANG F, DASHTIPOUR K, HUSSAIN A. A survey on deep learning for the routing layer of computer network[C]//Proceedings of 2019 UK/China Emerging Technologies (UCET). Piscataway: IEEE Press, 2019: 1-4.
- [24] LITTMAN M L. Markov games as a framework for multi-agent reinforcement learning[M]//Machine learning proceedings 1994. San Francisco: Morgan Kaufmann, 1994: 157-163.
- [25] VINYALS O, BABUSCHKIN I, CZARNECKI W M, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning[J]. Nature, 2019, 575(7782): 350-354.

[作者简介]



丛培壮(1994—), 男, 北京邮电大学网络与交换技术国家重点实验室博士生, 主要研究方向为网络体系结构、移动互联网创新服务应用等。



张宇超(1989—), 女, 北京邮电大学计算机学院(国家示范性软件学院)副教授、硕士生导师, 主要研究方向为数据驱动网络、大规模数据中心网络、边缘计算。

田野(1981—), 男, 北京邮电大学网络与交换技术国家重点实验室副教授、硕士生导师, 主要研究方向为下一代互联网、数据挖掘、知识发现等。

王文东(1963—), 男, 北京邮电大学网络与交换技术国家重点实验室教授、博士生导师, 主要研究方向为网络体系结构、网络和业务的服务质量控制与管理、机器学习与移动互联网网络与应用等。

李丹(1981—), 男, 博士, 清华大学教授、博士生导师, 主要研究方向为数据中心网络、网络智能和可信任互联网。入选教育部“青年长江学者”奖励计划, 获得国家优秀青年基金项目资助, 国家“973”项目首席科学家, 国家重点研发计划项目负责人。IEEE Transactions on Computers、IEEE Transactions on Wireless Communications 等国际学术期刊编委, ACM SIGCOMM、IEEE INFOCOM 等国际学术会议程序委员会委员。