



基于图神经网络的智能路由机制

张 鹏,陈 博

(中国人民解放军战略支援部队信息工程大学,郑州 450002)

摘 要: 现有基于人工智能的路由方案泛化能力较差,难以适应动态的网络拓扑变化。提出基于深度强化学习的智能路由机制 SmartRoute。通过实时感知网络中流量分布状态,动态调整路由策略,并结合图神经网络的拓扑信息感知能力和深度强化学习的自我训练能力,提升网络路由策略的智能性。实验结果表明,与 DRL-TE、TIDE 等方案相比,SmartRoute 最多节省 9.6% 的端到端时延,且具有更好的鲁棒性。

关键词: 软件定义网络;路由机制;深度强化学习;图神经网络;人工智能

开放科学(资源服务)标志码(OSID):



中文引用格式:张鹏,陈博.基于图神经网络的智能路由机制[J].计算机工程,2021,47(12):171-176,184.

英文引用格式:ZHANG P, CHEN B. Intelligent routing mechanism based on graph neural networks[J]. Computer Engineering, 2021, 47(12): 171-176, 184.

Intelligent Routing Mechanism Based on Graph Neural Networks

ZHANG Peng, CHEN Bo

(People's Liberation Army Strategic Support Force Information Engineering University, Zhengzhou 450002, China)

[Abstract] Existing artificial intelligence-based routing schemes are limited in generalization performance, and fail to adapt to the topological changes of networks. This paper proposes an intelligent routing strategy named SmartRoute based on deep reinforcement learning. SmartRoute can dynamically adjust the routing strategy by sensing the network traffic distribution in real time. Additionally, it combines the topology information perception ability of graph neural network and the self-training ability of deep reinforcement learning to improve the intelligence of network routing strategy. Experimental results show that SmartRoute saves up to 9.6% of end-to-end delay, and exhibits higher robustness than DRL-TE, TIDE and other schemes.

[Key words] Software-Defined Network (SDN); routing mechanism; deep reinforcement learning; Graph Neural Networks (GNN); artificial intelligence

DOI: 10.19678/j.issn.1000-3428.0059601

0 概述

近年来,随着信息技术与人类经济社会的深度融合,人们对于信息传输的需求不断增长。远程会议、5G、虚拟现实(Virtual Reality, VR)等新兴技术的蓬勃发展,极大增加了网络传输的信息量需求,给信息网络基础设施的传输能力带来了巨大的承载压力,也对网络服务质量的精细化提出了更高的要求。针对网络中不断增长的流量需求,传统的方法通常依赖于物理设备的扩容,包括增设传输线路、增加交换机数量等。然而,由于网络流量存在巨大的波动性,单纯的物理设备扩容面临着资源利用率低、更新周期长等问题,不能在有限成本下实现更好的网络

- 传输效果和网络服务质量(Quality of Service, QoS)^[1]。

通过网络流量的路由控制,能够使网络流量更加合理地分布于不同的网络设备上,从而减少网络资源使用的不均衡情况,提升网络的平均资源利用率。传统网络的路由算法基于不同路由器中运行的分布式路由算法实现。分布式路由算法通常按照“尽力而为”的原则进行路由计算,对于流量的控制精度有限,无法取得较好的网络服务质量。随着软件定义网络(Software-Defined Network, SDN)技术的出现,基于集中化网络视图的流量控制策略得以实现。SDN控制器通过收集全网信息,能够对基于网络状态的宏观掌控信息制定网络策略,从而为网络提供更加精细化的服务策略。

基金项目:国家重点研发计划“多模态智慧网络核心技术与原理平台”(2019YFB1802502)。

作者简介:张 鹏(1982—),男,副研究员,主研方向为网络架构、网络协议;陈 博(通信作者),讲师、博士研究生。

收稿日期:2020-09-27 修回日期:2020-12-12 E-mail: daniel_0720@163.com

然而,基于网路全局信息制定最优化的网络路由策略已经被证明为NP难问题^[2]。现有的路由优化算法为保证路由策略生成的动态性和实效性,通常基于人工设计启发式算法实现,无法保证路由策略的质量。针对该问题,业界将目光放在近年来发展迅速的人工智能技术。基于人工神经网络(Artificial Neural Network, ANN)、深度强化学习(Deep Reinforcement Learning, DRL)等技术的路由方案优势渐显。其中,基于深度强化学习的智能路由方案相比其他机器学习方案具有自主训练、适应性强、无需人工标记大量数据等优势,在众多基于机器学习的方案中脱颖而出。然而,现有的基于DRL实现的智能路由方案大多基于前馈神经网络(Feedforward Neural Network, FNN)、循环神经网络(Recurrent Neural Network, RNN)等神经网络结构实现,其网络结构固定、泛化能力差,难以适应动态的网络拓扑变化。

针对上述问题,本文提出一种基于图神经网络(Graph Neural Network, GNN)实现的智能路由机制SmartRoute。SmartRoute基于DRL框架实现GNN参数的智能调优,在训练环境下不断迭代更新策略参数,自动搜索网络的最优路由策略。

1 相关研究

当前基于机器学习算法实现的智能路由方案由于需要全局化网络视图信息,通常基于SDN框架实现。按照机器学习算法的特征,主要可以分为基于监督学习、基于无监督学习和基于强化学习的3类智能路由实现方案。

基于监督学习的方案主要通过深度神经网络(Deep Neural Network, DNN)实现。HUANG等^[3]使用DNN设计一种流量分类方案,基于不同流量类别分别适配不同的静态路由方案;文献[4-5]提出基于DNN的流量分类方案,通过不同流量类别实现不同的路由算法完成网络流量的调度。MAO等^[6]提出基于深度置信网络(Deep Belief Network, DBN)的智能路由方案,通过不同的流量特征实现网络边缘节点和核心节点的路由调整。总体来说,基于监督学习的智能路由方案需要人工标记大量网络流量特征,难以实现通用性。

基于无监督学习的方案通常也是基于网络流量特征分析实现网络路由策略。例如,文献[7-8]提出使用K-means算法实现网络流量的分类,文献[9]使用主成分分析(Principal Component Analysis, PCA)法实现网络特征提取。完成网络流量的特征分析后,上述方案仍然需要人工设计相应的路由策略,并且由于无监督学习的算法精度通常较难保证,因此基于无监督学习的智能路由方案通常难以取得较好的性能。

基于强化学习算法的智能路由方案通过强化学习算法(包括DRL算法)实现网络的智能路由。不同于前述2类智能路由方案,基于强化学习算法的智能路由方案能够直接实现从网络输入信息(例如流量分布视图)到路由策略的映射,并且能够在训练环境下实现自主调优,因此,其路由方案通常能够取得更优的性能。BOYAN等^[10]提出基于Q-learning算法的路由方案,能够智能调整路由的下一条选择;DRL-TE^[11]通过DRL算法调整多路径路由的不同路径分流比,优化了多路径路由的性能;文献[12-13]使用DRL算法调整网络拓扑中的路由权重,从而实现了全网路由计算结果的动态调整。总体来说,基于强化学习的智能路由方案已经成为当前智能化路由方案的研究主体。然而,在上述基于强化学习的智能路由方案中,神经网络主要采用前馈神经网络和循环神经网络等传统神经网络结构,此类神经网络结构只能处理欧式输入数据结构(例如固定维度的一维文本数据或者二维图像数据),对于输入数据格式具有严格的确定性限制;而在实际应用中,网络拓扑变化后神经网络输入数据格式也将发生变化,传统神经网络无法将训练经验泛化到新的网络拓扑中。因此,现有的基于强化学习实现的智能路由方案对于训练过程所使用的网络拓扑具有过拟合性,无法灵活应用于不同网络拓扑,也不能较好地处理网络链路故障等被动性网络拓部变化。

本文针对上述研究,提出一种基于图神经网络的智能路由机制。本文方案的主要贡献如下:

- 1)提出一种基于DRL实现的网络路由动态调整机制,能够随着网络流量的波动性实现流量的动态化调度。
- 2)设计一种结合GNN与DRL技术的智能训练算法,能够大幅提高智能路由算法的泛化能力,提升算法的鲁棒性。
- 3)设计基于NS2的仿真环境,完成智能路由策略的离线训练。

2 方案设计

本节主要介绍了结合GNN和DRL的智能路由方案SmartRoute,阐述了SmartRoute的主要架构和各部分实现机制。

2.1 SmartRoute架构

SmartRoute主要基于SDN框架实现,其架构如图1所示。其中,实现智能路由的功能主要由数据平面的可编程交换机、控制平面的SDN控制器以及控制器之上运行的DRL算法实现。可编程交换机负责统计网络中的流量信息(包括实时流量分布、流量传输性能等指标);SDN控制器负责收集和统计流

量信息、更新交换机的转发表;DRL算法负责根据网络中的状态信息生成输出动作,该输出动作值直接用于路由策略生成。

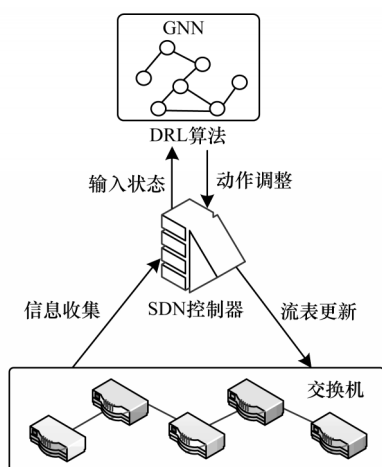


图1 SmartRoute 架构

Fig.1 SmartRoute architecture

DRL算法结合图神经网络进行网络输入数据的动态分析,将生成的输出动态作为链路权重,基于此动态调整链路权重,SDN控制器通过加权最短路径算法可以灵活调整网络路由,从而实现路由的智能控制。其中,由于SmartRoute基于图神经网络进行网络流量数据的分析,在网络拓扑变化时,Smart Route具有良好的泛化性能,从而提升不同网络环境下路由性能的稳定性。

2.2 DRL 算法框架

DRL算法由强化学习的基本框架扩展而来。强化学习通过强化学习算法与其运行环境之间的信息交互,不断优化算法参数,从而实现算法的训练。其中,强化学习的基本模型将网络建模为一个马尔科夫过程,其算法的主要交互元素包含 (s_t, a_t, r_t) 3个部分, s_t 表示 t 时刻下环境的状态信息, a_t 表示 t 时刻下算法的输出动作, r_t 表示对于 t 时刻算法动作内容产生结果的回报值。由状态信息生成输出动作可以通过查表、函数计算等不同方式实现,而深度强化学习正是使用了深度神经网络完成这一过程,因此被称为深度强化学习。深度强化学习的学习目标是在一个阶段内实现回报值的最大化积累,即 $\max \sum_{t=0}^T r_t$,或者添加折扣因子 γ 的最大化折扣回报值累积 $\max \sum_{t=0}^T \gamma^t r_t$ 。

深度强化学习通过奖励值 r_t 实现神经网络参数的调整,根据具体实现方式的不同可以分为值优化、策略梯度优化等多种算法框架。本文所使用的深度强化学习框架为策略梯度优化。策略梯度优化方法可以通过式(1)表述:

$$\nabla_{\theta} E_{\pi} \left[\sum_{k=0}^T r_k \right] = E_{\pi} \left[\sum_{k=0}^T \nabla_{\theta} \log_{\pi} (s_k, a_k) Q(s_k, a_k) \right] \quad (1)$$

其中: $Q(s_k, a_k)$ 为当前状态-动作内容的评估价值; π 表示当前的动作生成策略。在每个时刻,由式(1)可以计算出神经网络的更新梯度值。在实际更新神经网络参数的过程中,SmartRoute采用文献[14]的实现方案,即可以通过对于 Q 值的变形处理实现神经网络的更新梯度值计算,如式(2)所示:

$$\theta \leftarrow \theta + \alpha \sum_{k=0}^T \nabla_{\theta} \log_{\pi} (s_k, a_k) v_k \quad (2)$$

其中: θ 表示神经网络的参数集合; v_k 由折扣回报值计算得到,作为梯度更新值的加权因子。SmartRoute的整体DRL算法如下:

算法1 DRL 算法

1. 随机初始化神经网络参数 θ ;
2. For $i = 1$ to M :
3. 运行交互过程 $(s_1^i, a_1^i, r_1^i, \dots, s_L^i, a_L^i, r_L^i) \sim \pi_{\theta}$;
4. 计算累积回报值 $R_L^i = \sum_{t'=1}^L r_{t'}^i$;
5. $\Delta\theta = 0$;
6. For $t = 1$ to L :
7. $b_t = \frac{1}{N} \sum_{i=1}^N R_t^i$;
8. For $i = 1$ to N :
9. $\Delta\theta = \Delta\theta + \nabla_{\theta} \log \mu_{\theta}(s_t^i, a_t^i) (R_t^i - b_t)$;
10. End for
11. End for
12. $L = L + \epsilon$;
13. $\theta = \theta + \alpha \Delta\theta$;
14. End for

2.3 神经网络结构设计

网络中流量分布等信息的呈现结构依赖于网络拓扑形状。为更好地处理网络拓扑结构信息,提升神经网络对于不同拓扑结构数据的泛化能力,本文使用GNN作为DRL算法框架的神经网络实现。具体来讲,本文所使用的GNN类型为信息传递神经网络(Message Passing Neural Network, MPNN)^[15]。MPNN的主要功能通过信息传递过程实现,其过程如图2表示。图2左侧展示一个包含4个图节点的MPNN,其中箭头表示节点4接收其他节点传输信息的过程, h_i 表示节点 i 的隐藏状态。在每一时刻,每个节点按照图2所示的方式与其所有邻居节点交换信息,交换信息的计算方式由图2中的函数(神经网络) $m(\cdot)$ 和 $u(\cdot)$ 实现。通过每个节点信息的不断迭代,MPNN最终完成网络的整体信息更新。由于SmartRoute中的MPNN主要处理流量变化信息,因此在实现过程中,函数 $m(\cdot)$ 由全连接神经网络实现, $u(\cdot)$ 通过门控循环单元(Gated Recurrent Unit, GRU)实现。

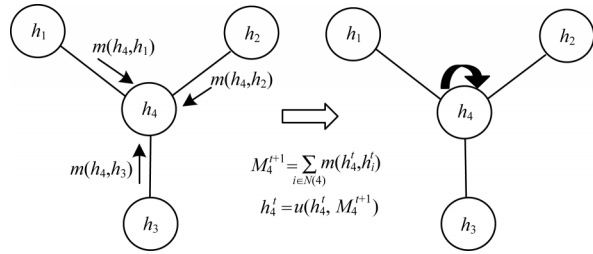


图2 MPNN工作流程

Fig.2 The working procedure of MPNN

在 SmartRoute 中,实际的信息通信网络与 MPNN 中的图采用不同建模方式进行映射。具体来讲,MPNN 主要通过其图中节点信息的交互完成网络的信息传输和计算,而 SmartRoute 通过调整链路权重完成网络路由的调整,因此其调整的元素为网络拓扑中的链路信息。为使用 MPNN 实现信息通信网络中链路权重的调整,SmartRoute 将通信网络的链路映射为 MPNN 的节点,通过 MPNN 中节点信息的传递和计算,最终完成信息通信网络中链路的信息更新和计算。其中,信息通信网络中链路信息包括链路的实时带宽占用信息。整体的计算过程如图 3 所示。

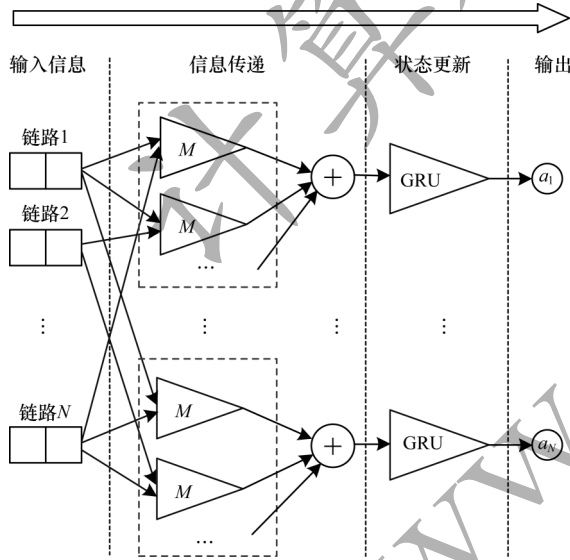


图3 网络信息计算过程

Fig.3 Calculation process of network information

基于上述神经网络设计,SmartRoute 中 DRL 的基本接口可以表述如下:

1) 状态。SmartRoute 中 DRL 算法的状态信息为链路利用率和链路时延信息。每条链路的上述信息按照一定的格式输入,作为 MPNN 对应节点的输入信息。

2) 动作。SmartRoute 中 DRL 算法的动作内容为链路权重。MPNN 经过计算后,每个节点输出值对应于网络中的所有链路权重信息。SDN 控制器根据

此链路权重信息可以计算网络的端到端加权最短路径。

3) 回报值。回报值的设计决定了 DRL 算法的自主优化方向,也即网络性能的优化方向。常见的性能指标可以结合网络的总体吞吐率、平均端到端传输时延等性能指标设计。SmartRoute 中采用网络的平均端到端时延作为回报值。

在每一步 DRL 交互过程中,MPNN 的计算过程伪代码如表 2 所示。

算法2 MPNN算法

输入 链路信息 $[h_1, h_2, \dots, h_N]$

输出 动作值 $[a_1, a_2, \dots, a_N]$

1. For $t=1$ to T :

2. For $l=1$ to N :

3. $M_l^{t+1} = \sum_{i \in N(l)} m(h_l^t, h_i^t)$

4. $h_l^{t+1} = u(h_l^t, M_l^{t+1})$

5. End for

6. End for

7. For $l=1$ to N :

8. $a_l \leftarrow NN(h_l)$

9. End for

3 仿真结果与评估

本节主要阐述 SmartRoute 的仿真验证环境以及相关仿真结果的分析。

3.1 仿真环境及参数设计

本文基于 NS2 环境实现了 SmartRoute 的仿真验证平台。在 SmartRoute 中所使用的 DRL 算法基于 Tensorflow 1.12 实现,采用 Python3.7 语言编写。在 Tensorflow 与 NS2 之间,本文采用 Python 语言编写了基于 OpenAI 的 Gym 环境^[16]接口用于连接 DRL 算法和训练环境。实验的硬件实现平台具有一颗 i7 9700K CPU、16 GB DDR4 内存和一块 1080Ti 显卡。

SmartRoute 中 DRL 算法迭代次数设置为 60 000 次,算法 2 中 MPNN 的迭代次数 T 设置为 8 次,神经网络参数优化方式采用基于 Nesterov Momentum 的随机梯度下降^[17]法,学习率设置为 5×10^{-5} 。在本文的硬件实验平台下,完成一次算法训练耗时约为 6 h。训练完成后,所得方案可以部署到不同网络拓扑中,实现快速路由计算(OS3E 拓扑下 SmartRoute 每次计算生成新的路由规则耗时 287 μ s)。

在仿真环境中,采用了 Topology Zoo^[18]中选取的 OS3E 拓扑,每个拓扑的链路带宽容量统一设置为 100 Mb/s。网络中的流量生成按照随机流量与周期性流量组合的方式生成^[12,19-21],通过调整随机流量与周期性流量的比例改变流量特征。每个网络节点都可以与任意一个其他网络节点发起通信需求,其平均流量设置为网络总吞吐量的特定比例(20%、30%、40%)。

3.2 对比方案

在本文实验中,SmartRoute 主要与以下3种方案进行性能对比:

1)TIDE^[12]。TIDE 通过使用DDPG算法框架,用循环神经网络作为主要神经网络对网络链路的流量分布进行计算,输出不同链路的权重信息。通过更新链路权重信息,网络可以更新路由计算结果。

2)DRL-TE^[11]。将网络端到端通信通过多路径实现(每个端到端通信设置3条备份路径),通过DRL算法感知网络中的流量分布,调节每个数据流在3条备份路径上的分流比,从而实现流量工程。

3)ECMP。ECMP通过等价多路径传输网络中的流量,为网络中的典型流量工程方案之一。

3.3 性能评估

性能评估主要包括以下3个方面:

1)端到端时延对比。本文实验主要使用网络中的平均端到端时延作为网络性能的衡量指标。分别设置了3组不同的网络流量特征(即不同的周期流量/随机流量)。图4~图6分别展示了不同流量强度下OS3E拓扑中流量特征的平均端到端时延。由图4~图6的总体趋势可以看出,随着网络中总流量的提高,各个方案的平均端到端时延都呈现出上升趋势,符合直观推断。图4~图6中横坐标表示由不同的流量构成,以图4为例(图5、图6具有相同规律),可以看出,随着随机流量的增多,DRL-TE、TIDE和SmartRoute的平均端到端时延呈现出了明显增长的趋势,而ECMP的平均端到端时延具有相对稳定的数值,只有流量的时延方差明显增大。这是因为DRL-TE、TIDE和SmartRoute都是通过训练后的神经网络提取网络中的流量规律从而制定相应的路由策略,而随着随机流量的增多,流量所呈现出的规律性明显减弱,上述3种方案对于流量特征提取的能力也相应减弱,造成路由策略的时延性能下降。总体来讲,SmartRoute在不同流量特征下的时延大多优于其他方案,在最佳情况下能够比当前最优方案节省9.6%的平均端到端时延。

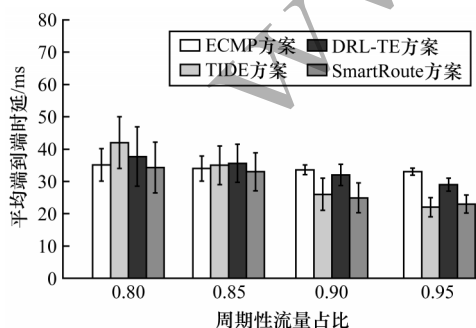


图4 20%吞吐量下平均端到端时延

Fig.4 Average end-to-end delay under the 20% throughput

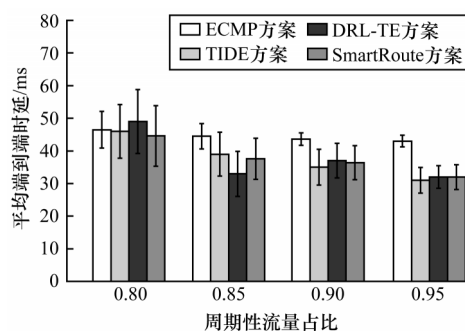


图5 30%吞吐量下平均端到端时延

Fig.5 Average end-to-end delay under the 30% throughput

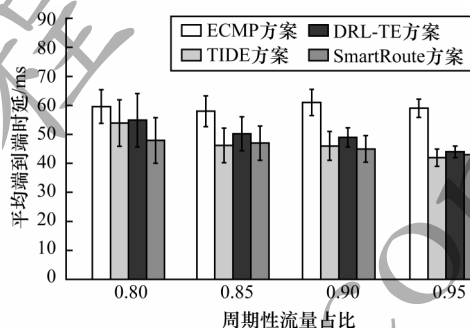


图6 40%吞吐量下平均端到端时延

Fig.6 Average end-to-end delay under the 40% throughput

2)路由策略的泛化能力。常见的人工神经网络通常具有过拟合的缺点,即在训练过程中的输入数据格式需要保持不变,而训练完成后,该神经网络也难以适应不同的输入数据格式。DRL算法主要依赖于神经网络计算产生策略,因此采用循环神经网络等神经网络的DRL算法也存在过拟合的问题。在本文实验场景中,在OS3E拓扑下完成不同的路由方案训练后,将所得路由算法应用到其他网络,以测试不同方案的泛化能力,其结果如图7所示。其中,横坐标表示以OS3E拓扑为基础不同测试拓扑的节点数量变化(即增多的节点数量),纵坐标表示相对于OS3E原始拓扑下网络的端到端时延变化(正值表示时延性能有优化,负值表示时延增加)。

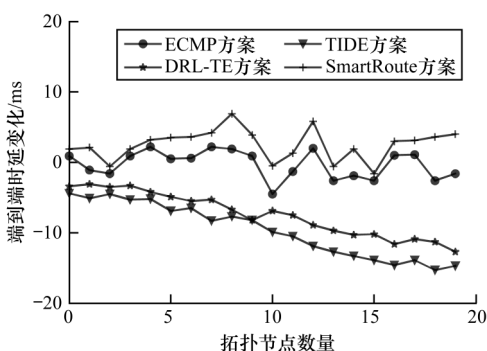


图7 不同路由方案泛化能力

Fig.7 The generalization ability of different routing schemes

由图7数据可以看出,随着网络拓扑变化越来越大(比原始训练拓扑OS3E增加节点越来越多),TIDE和DRL-TE的流量传输呈现出明显的时延增加趋势,而ECMP和SmartRoute则无明显的性能下降趋势。其中,ECMP和DRL-TE在不同拓扑下的时延波动性主要由拓扑本身的性质造成。

3)鲁棒性。本文实验主要测试SmartRoute等方案在网络链路失效情况下的路由鲁棒性,从而展示其在实际网络运行环境下的可用性能。在测试中,将OS3E拓扑随机选取不同数量的链路断开(若所选链路破坏了拓扑的连通性,则重新选择链路),以模拟真实网络运行环境下的网络链路故障。由于TIDE和DRL-TE依赖于原始网络拓扑,无法适应链路故障,因此本测试仅对比ECMP和SmartRoute的链路故障后路由性能。从图8的实验结果可以看出,在链路失效后,SmartRoute总是能够比ECMP具有更低的端到端时延,证明了SmartRoute方案在网络链路故障场景下路由性能的鲁棒性。

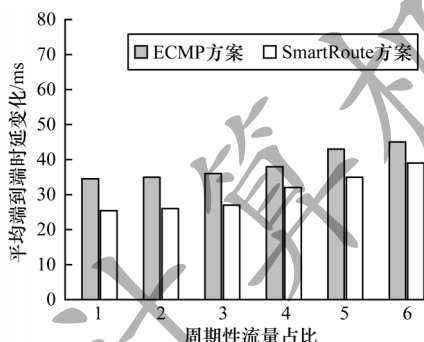


图8 链路故障下的鲁棒性测试结果

Fig.8 Robustness test results under link failure

4 结束语

本文结合GNN和DRL技术,提出一种智能路由机制SmartRoute。SmartRoute使用DRL算法实现方案参数的自主训练,网络的交互过程不依赖于人工经验搜索最优策略。同时,通过使用GNN计算网络中基于拓扑结构的流量分布数据,能够实现在不同网络拓扑下路由策略的泛化能力和路由策略的鲁棒性。下一步将通过构建深度学习模型,对神经网络的设计结构进行优化,提升网络的智能路由策略性能。

参考文献

- [1] PAXSON V, STANIFORD S. The cooperative associate for Internet data analysis[EB/OL]. [2020-08-11]. <http://www.caida.org/data>.
- [2] 崔勇, 吴建平, 徐恪. 互联网服务质量路由算法研究综述[J]. 软件学报, 2002, 13(11): 2065-2075.
CUI Y, WU J P, XU K. A survey on quality-of-service routing algorithms for the Internet[J]. Journal of Software, 2002, 13(11): 2065-2075. (in Chinese)
- [3] HUANG W H, SONG G J, HONG H K, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2191-2201.
- [4] KATO N, FADLULLAH Z M, MAO B, et al. The deep learning vision for hetero-geneous network traffic control: proposal, challenges, and future perspective [J]. IEEE Wireless Communications, 2017, 24(3): 146-153.
- [5] 刘佳美, 徐巧枝. 基于机器学习的SDN网络流量预测与部署策略[J]. 计算机工程, 2020, 46(10): 223-230.
LIU J M, XU Q Z. Network traffic prediction and deployment strategy based on machine learning for SDN[J]. Computer Engineering, 2020, 46(10): 223-230. (in Chinese)
- [6] MAO B M, FADLULLAH Z M, TANG F X, et al. Routing or computing? The paradigm shift towards intelligent computer network packet transmission based on deep learning [J]. IEEE Transactions on Computers, 2017, 66(11): 1946-1960.
- [7] ERMAN J, ARLITT M, MAHANTI A. Traffic classification using clustering algorithms [C]//Proceedings of 2006 SIGCOMM Workshop on Mining Network Data. New York, USA: ACM Press, 2006: 281-286.
- [8] LIU Y Q, LI W, LI Y C. Network traffic classification using k-means clustering [C]//Proceedings of International Multi-Symposiums on Computer and Computational Sciences. Washington D. C., USA: IEEE Press, 2007: 360-365.
- [9] YAN R Y, LIU R. Principal component analysis-based network traffic classification [J]. Journal of Computers, 2014, 9(5): 1234-1240.
- [10] BOYAN J A, LITTMAN M L. Packet routing in dynamically changing networks: a reinforcement learning approach [C]//Proceedings of Advances in Neural Information Processing Systems. San Francisco, USA: Morgan Kaufmann Press, 1994: 671-678.
- [11] XU Z Y, TANG J, MENG J S, et al. Experience-driven networking: a deep reinforcement learning based approach [C]//Proceedings of IEEE Conference on Computer Communications. Hawaii, USA: IEEE Press, 2018: 1871-1879.
- [12] SUN P H, HU Y X, LAN J L, et al. TIDE: timerelevant deep reinforcement learning for routing optimization [J]. Future Generation Computer Systems, 2019, 11(3): 134-144.
- [13] SUN P H, LI J F, GUO Z H, et al. SINET: enabling scalable network routing with deep reinforcement learning on partial nodes [C]//Proceedings of SIGCOMM'19. Beijing, China: [s. n.], 2019: 88-89.
- [14] RONALD J. Simple statistical gradient-following algorithms for connectionist reinforcement learning [J]. Machine Learning, 1992, 3(4): 229-256.
- [15] GILMER J, SCHOENHOLZ S S, RILEY O F, et al. Neural message passing for quantum chemistry [C]//Proceedings of the 34th International Conference on Machine Learning. New York, USA: ACM Press, 2017: 1263-1272.

(下转第184页)

(上接第 176 页)

- [16] BROCKMAN G, CHEUNG V, PETTERSSON L, et al. OpenAI GYM[EB/OL]. [2020-08-10]. <https://arxiv.org/abs/1606.01540>.
- [17] BOTTOU L. Large-scale machine learning with stochastic gradient descent[C]//Proceedings of COMPSTAT' 10. Berlin, Germany: Springer, 2010: 177-186.
- [18] SIMON K, HUNG X, NICKOLAS F, et al. The Internet topology zoo [J]. IEEE Journal on Selected Areas in Communications, 2011, 29(9): 1765-1775.
- [19] LAKHINA A, PAPAGIANNAKI K, CROVELLA M, et al. Structural analysis of network traffic flows[C]//Proceedings of Joint International Conference on Measurement and Modeling of Computer Systems. New York, USA: ACM Press, 2004: 61-72.
- [20] GUO Z H. Joint switch upgrade and controller deployment in hybrid software-defined networks[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(5): 1012-1028.
- [21] GUO Z H. RetroFlow: maintaining control resiliency and flow programmability for software-defined WANs[C]//Proceedings of the 27th IEEE/ACM International Symposium on Quality of Service. Phoenix, USA: ACM Press, 2019: 1-14.

编辑 索书志