



面向确定性网络的按需智能路由技术

伍仲丽¹, 曹园园¹, 黄文睿¹, 戴彬¹, 莫益军²

(1. 华中科技大学电子信息与通信学院, 湖北 武汉 430074;

2. 华中科技大学计算机科学与技术学院, 湖北 武汉 430074)

摘要: 确定性网络需要保证不同应用在时延、丢包率、抖动、吞吐量和可靠性等方面的确定性传输需求。针对应用的差异化、确定性的网络传输需求, 提出了一种面向确定性网络的按需智能路由学习框架 OdR, 在 OdR 框架下提出一种基于深度强化学习的按需智能路由算法 OdR-TD3, OdR-TD3 算法可以根据应用流量的确定性 QoS 需求生成路由策略, 以满足确定性网络应用的需求。通过网络仿真实验评估, 在确定性应用的 QoS 需求达成率上, OdR-TD3 算法相较 DV 算法和 SPF 算法, 具有显著的优势。

关键词: 确定性网络; 深度强化学习; 按需智能路由; 服务质量

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021245

On-demand intelligent routing technology for deterministic network

WU Zhongli¹, CAO Yuanyuan¹, HUANG Wenrui¹, DAI Bin¹, MO Yijun²

1. School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China

2. School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China

Abstract: Deterministic network needs to ensure the deterministic transmission requirements of different applications in terms of delay, packet loss rate, jitter, throughput, and reliability. In response to the differentiated and deterministic network transmission requirements of applications, an on-demand intelligent routing framework OdR for deterministic network was proposed. Under the OdR framework, an on-demand intelligent routing algorithm named OdR-TD3 based on deep reinforcement learning was proposed, which generates routing strategies based on the deterministic QoS requirements of application traffic, to satisfy the applications' requirements of deterministic network. The experimental evaluation results show the OdR-TD3 algorithm has a significant advantage over the DV algorithm and the SPF algorithm in terms of the achievement rate of deterministic QoS requirements.

Key words: deterministic network, deep reinforcement learning, on-demand intelligent routing, quality of service



1 引言

随着新型网络应用的不断涌现,如 VR/AR、全息通信、触觉互联网、工业互联网、智能电网、自动驾驶、远程医疗等,新型应用对网络传输提出了新的需求和挑战。不同应用在带宽速率、时延、抖动、能耗、服务安全等方面有着差异化的需求^[1],而且对于时延、抖动、丢包等关键服务质量 (quality of service, QoS) 指标提出了确定性 (有上下界) 的要求。网络传输需求的差异化和确定性的 QoS 需求给基础网络运维带来了极大的挑战^[2]。

在确定性网络研究方面,时间敏感网络 (time-sensitive network, TSN) 和确定性网络 (deterministic network, DetNet) 工作组开展了颇有成效的工作^[3],致力于为应用流量提供最优的路径,以满足应用对时延、抖动、丢包等的确定性 QoS 需求。但网络环境的复杂性和不确定性,给传统的数学建模求解带来了极大的困难^[4]。近年来,有学者将人工智能技术应用于网络路由优化,借助机器学习能够摆脱精确的数学模型,可以处理更加复杂的网络环境,基于强化学习的路由决策能够实时响应环境和用户的需求变化,已成为网络路由优化的热点研究方向^[5]。

网络应用对于时延、抖动、丢包等关键性能指标提出了确定性 QoS 保障的需求。例如对于 VR/AR 应用,数据率要求为 40 Mbit/s (2D)/63 Mbit/s (3D),时延要求低于 30 ms (2D) / 20 ms (3D),分组丢失率要求小于 2.40×10^{-5} 。对于工业互联网应高水平,同时网络边缘的其他路由节点则常处于闲置状态,导致网络整体性能下降,数据率要求 1~10 Gbit/s,时延要求低于 5 ms。对于远程医疗应用,数据率要求为 100 Mbit/s,时延要求低于 100 ms。在网络结构和应用流量不断变化的复杂环境中,如何满足不同应用的确定性 QoS 需求成为网络路由优化的难点。

传统路由算法有距离矢量 (distance vector, DV) 算法^[6]、最短路径优先 (shortest path first, SPF) 算法^[7]和启发式算法等,其中, DV 算法收敛速度较慢,并且在收敛过程中可能会遇到路由选择环路,产生无穷计数问题。SPF 算法大多以跳数为度量,是一种尽力而为的路由方案,选择固定的路径,这些选定路径上的节点负载始终维持在较高水平,导致负载不均衡,降低了网络整体性能。启发式算法的收敛速度慢、训练时间长、计算复杂度高,因此不适合越来越复杂的网络架构。这些路由算法没有充分利用丰富的网络资源,只能提供非常有限的 QoS 保证,并且传统路由算法没有时序和记忆功能,往往不会从以往的路径拥塞、严重延迟甚至丢包的经历中吸取教训。与传统路由算法相比,机器学习技术具有一定的优越性,经过训练,机器学习算法面对给定的输入、输出接近最优的路由策略,其次,机器学习不依托精确的底层网络数学模型,另外,机器学习在分析和处理大量数据方面具有强大能力,能发现复杂环境的数据模式并做出准确的决策。最后,机器学习还可以从过去无效的路由决策和拥塞、丢包等事件中学习教训,避免类似情况再次出现^[8]。深度强化学习 (deep reinforcement learning, DRL) 结合深度学习和强化学习的优点,在训练阶段主动了解不断变化的网络状态,并在复杂的环境中针对不同的服务请求智能地选择合适的路径。但现有的智能路由优化方法,大多以网络整体效用最大化作为优化目标,未考虑到应用对确定性 QoS 指标的差异化需求^[9]。本文的研究针对网络应用的确定性 QoS 需求,基于深度强化学习算法的设计按需智能路由策略,满足应用的差异化、确定性的 QoS 需求。

2 模型与算法设计

2.1 OdR 模型

基于软件定义网络 (software defined network,

SDN) 的架构^[10], 本文提出了一种面向确定性网络的按需智能路由学习框架 OdR (on-demand intelligent routing framework), 如图 1 所示。OdR 模型包括 3 个逻辑平面: 智能路由平面、网络测量平面和数据转发平面。

- 智能路由平面是 OdR 模型的控制中枢, 多种具有不同确定性 QoS 需求的应用流量信息作为输入, 输出为各类确定性 QoS 需求所对应的链路权重, 通过 Dijkstra 算法将链路权重转换为路由策略, 生成的路由策略发布到数据转发平面。
- 网络测量平面负责收集网络状态 (如流量需求、资源可用性、等待时间等) 和应用流量的 QoS 指标, 用于评估路由策略的性能。
- 数据转发平面根据应用所携带的确定性 QoS 需求来分类和调度网络中的流量, 并且支持在线更新路由策略。

2.2 OdR-TD3 算法

基于 OdR 模型, 本文设计了一种基于 TD3 算法 (twin delayed deep deterministic policy gradient algorithm)^[11] 的 QoS 自适应按需路由算法: OdR-TD3。OdR-TD3 通过 3 个信号 (状态、动作和奖励) 与网络环境进行交互。环境与主体之间的交互过程可视为马尔可夫决策过程 $M = (S, A, R, P, \gamma)$, 其中, S 是状态空间, A 是动作空间, R 是奖励空间, P 是转移概率函数 $p(s_{t+1}, r | s_t, a_t) \in P$, 折现因子 $\gamma \in [0, 1]$ 。代理在状态 s 下选择动作 a 在确定性策略中表示为 $a = \mu(s)$ 。

为了满足网络应用的确定性 QoS 需求, 奖励函数的设计至关重要, 设计了一种基于 QoS 感知的奖励函数。用来评价路由策略优劣的 QoS 指标有时延、抖动和丢包等, 奖励函数 R_t 如式 (1) 所示, 用于指导 OdR-TD3 算法的训

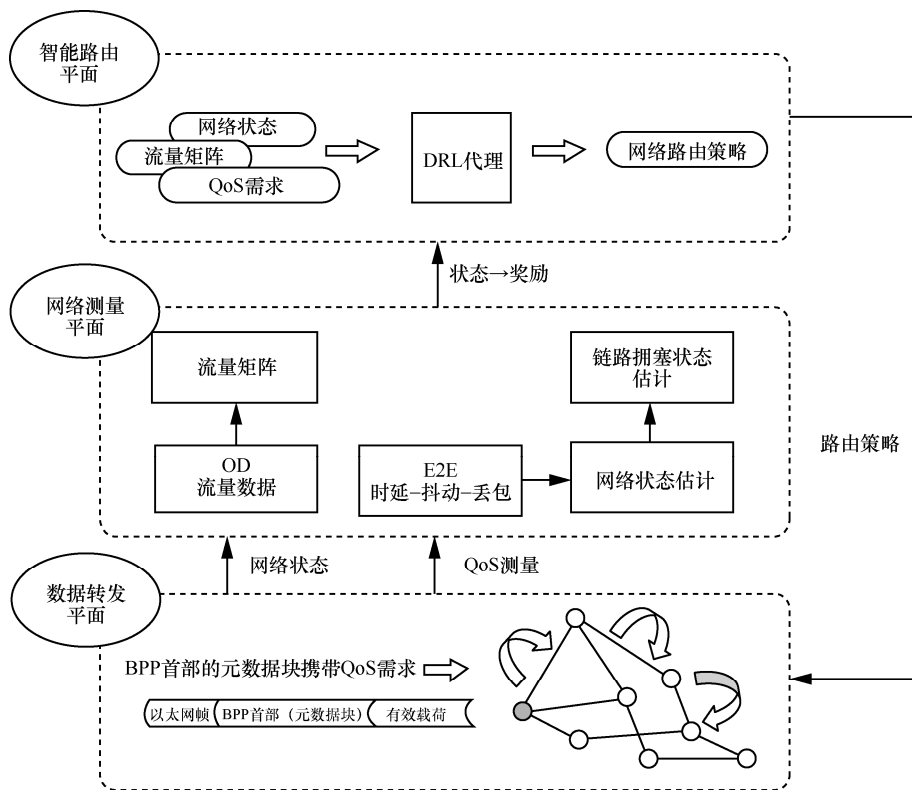


图 1 OdR 模型



练,目的是确定最优策略 π 关于状态空间 S 到动作空间 A 的映射 ($\pi: S \rightarrow A$), 最大限度地提高预期奖励 R_t 。

$$R_t = \sum_{i,j \leq n_y} R(i \rightarrow j | s_t, a_t) = - \sum_{k=1}^K \sum_{i,j \leq n_y} (\alpha_k \hat{D}_{i \rightarrow j}^k + \beta_k \hat{J}_{i \rightarrow j}^k + \gamma_k \hat{L}_{i \rightarrow j}^k + f_k(D_{i \rightarrow j}^k) + f_k(J_{i \rightarrow j}^k) + f_k(L_{i \rightarrow j}^k)) \quad (1)$$

其中, $D_{i \rightarrow j}^k, J_{i \rightarrow j}^k, L_{i \rightarrow j}^k$ 是源节点 i 到目的节点 j 的第 k 种应用类型的时延、抖动和丢包的度量, $\hat{D}_{i \rightarrow j}^k, \hat{J}_{i \rightarrow j}^k, \hat{L}_{i \rightarrow j}^k$ 是 $D_{i \rightarrow j}^k, J_{i \rightarrow j}^k, L_{i \rightarrow j}^k$ 的归一化量, 可调参数 $\alpha_k, \beta_k, \gamma_k \in [0,1]$ 表征相应的 QoS 指标对于第 k 种应用类型的重要程度。奖励函数中的 $f_k(x)$ 是一个惩罚函数, 在测量的 QoS 指标超过设定的界限情况下, 它会在奖励中引入一个额外的惩罚项。

OdR-TD3 算法过程见算法 1。

算法 1 OdR-TD3 算法

输入 K 种应用类型的流量需求

输出 K 种应用类型的路由策略

(1) 初始化 OdR-TD3 的 Critic 网络 Q_{ϕ_1}, Q_{ϕ_2} 和 Actor 网络 π_{θ} 与随机参数 ϕ_1, ϕ_2, θ , 目标网络参数 $\phi'_1 \leftarrow \phi_1, \phi'_2 \leftarrow \phi_2, \theta' \leftarrow \theta$ 和经验回放池 D

(2) 执行 M 回合语句 3~12

(3) 获取初始 K 种应用类型的流量需求 $\{T_t^1, T_t^2, \dots, T_t^K\}$ 作为 OdR-TD3 算法的输入

(4) 执行 T 次语句 5~12

(5) OdR-TD3 算法选择动作 $a \sim \pi_{\theta}(s) + \varepsilon$, 其中 $\varepsilon \sim N(0, \sigma)$, 动作经 softmax 函数处理后作为 K 种应用类型的链路权重 $\{W_t^1, W_t^2, \dots, W_t^K\}$

(6) 根据链路权重 $\{W_t^1, W_t^2, \dots, W_t^K\}$ 执行 Dijkstra 算法以计算 K 种应用类型的路由策略 $\{P_t^1, P_t^2, \dots, P_t^K\}$, 向数据转发平面发布路由策略, 转发节点根据数据包携带的 QoS 信息匹配对应类别的路由策略执行转发操作, 网络测量平面从全局监测网络状态

(7) 根据式 (1) 计算奖励 r , 获取下一时刻状态 s' , 将 $(s, a, r, s', \text{done})$ 存入经验回放池 D

(8) 从经验回放池 D 中随机小批量抽取 N 个样本 $(s_i, a_i, r_i, s_{i+1}, \text{done})$, $\tilde{a} \leftarrow \pi_{\theta'}(s') + \varepsilon$, $\varepsilon \sim \text{clip}(N(0, \tilde{\sigma}), -c, c)$, $y \leftarrow r + \gamma(1-d) \min_{i=1,2} Q_{\phi_i}(s', \tilde{a})$

(9) 更新 Critic 网络 $\phi_i \leftarrow \arg \min_{\phi_i} N^{-1} \sum (y - Q_{\phi_i}(s', \tilde{a}))^2$

(10) 每 d 步执行语句 (11) ~ (12)

(11) 根据确定性策略梯度 $\nabla_{\theta} J(\theta) = N^{-1} \sum \nabla_a \cdot Q_{\phi_1}(s, a)|_{a=\pi_{\theta}(s)} \nabla_{\theta} \pi_{\theta}(s)$ 更新 Actor 网络

(12) 更新目标网络: $\phi'_i \leftarrow \tau \phi_i + (1-\tau) \phi'_i$, $\theta'_j \leftarrow \tau \theta_j + (1-\tau) \theta'_j$,

(13) 保存训练好的 OdR-TD3 网络模型

3 实验分析

3.1 实验环境配置

实验网络拓扑采用了 23 个节点和 37 条全双工链路的 GÉANT。在 Ubuntu18.04 操作系统下, 使用 Keras 和 Tensorflow 实现深度强化学习算法, OMNeT++ 模拟网络环境。应用流量数据集来源于网上公开的网络流量跟踪数据集^[12]: 从 2005 年 1 月 8 日起每 15 min 汇总一次 GÉANT 的流量矩阵, 持续 16 周。实验都假设任意源节点不会向自身发送数据, 数据包到达每个源节点 (即流量需求) 遵循泊松过程。为了简化实验过程, 仅考虑网络中存在 3 类确定性 QoS 需求的应用, 3 类应用流量 $\alpha_1, \alpha_2, \alpha_3$ 的 QoS 指标需求见表 1。

表 1 3 类应用流量的 QoS 指标需求

应用流量	最大时延容忍	最大抖动容忍	最大丢包率容忍
α_1	100 ms	50 ms	0.01
α_2	200 ms	100 ms	0.02
α_3	1 s	300 ms	0.05

3.2 实验结果分析

对 OdR-TD3 算法和两种经典的路由算法 DV

算法和 SPF 算法进行了性能对比。在模型训练结束后保存 3 种算法的模型, 然后使用 200 组流量矩阵作为测试集分别对保存的模型进行测试, GÉANT 测试阶段平均奖励值见表 2, OdR-TD3 算法的平均奖励值最高。

表 2 GÉANT 测试阶段平均奖励值

	DV	SPF	OdR-TD3
奖励值	-2.335	-1.052	-0.538

GÉANT 测试阶段各类应用流量的 QoS 指标见表 3。从实验测试数据上可以看出, 在 QoS 的各项指标上, OdR-TD3 算法均大幅领先于 DV 算法和 SPF 算法, 其中平均时延较 DV 算法和 SPF 算法分别降低了 50%和 30%以上, 平均抖动则相较 DV 算法和 SPF 算法分别降低了 35%和 15%以上, 而平均丢包率相较 DV 算法和 SPF 算法分别下降了 80%和 60%以上。对比各类应用的确性 QoS 指标要求, 基于 OdR-TD3 算法的应用流量均能满足确定性 QoS 指标的要求, 且确定性 QoS 指标达成率远优于 DV 算法和 SPF 算法。

表 3 GÉANT 测试阶段各类应用流量的 QoS 指标

QoS 指标	应用类型	DV	SPF	OdR-TD3
平均时延/s	α_1	1.659	0.798	0.083
	α_2	1.754	0.857	0.157
	α_3	1.915	0.919	0.659
平均抖动/s	α_1	0.121	0.079	0.045
	α_2	0.134	0.085	0.076
	α_3	0.109	0.083	0.079
平均丢包率	α_1	0.057	0.026	0.005
	α_2	0.059	0.022	0.007
	α_3	0.067	0.019	0.006

4 结束语

本文主要研究了如何在确定性网络中保障各

类应用的确性 QoS 需求, 将 DRL 方法引入确定性网络路由中, 以实现按需的自适应智能路由决策。基于真实的数据集, 在特定的网络结构下对 OdR-TD3 算法进行评估, 与 DV 和 SPF 路由算法相比, OdR-TD3 算法表现出优异的性能, 可以更好地保障确定性的 QoS 需求。在网络仿真环节, 本文假设了 3 类确定性 QoS 需求的应用及其指标需求, 下一步研究中可以面向实际的确性应用对需求指标进行完善, 还可以验证算法在不同网络拓扑下的泛化性能。

参考文献:

- [1] KARAKUS M, DURRESI A. Quality of service (QoS) in software defined networking (SDN): a survey[J]. Journal of Network and Computer Applications, 2017(80): 200-218.
- [2] 黄韬, 汪硕, 黄玉栋, 等. 确定性网络研究综述[J]. 通信学报, 2019, 40(6):160-176.
HUANG T, WANG S, HUANG Y D, et al. Survey of the deterministic network[J]. Journal on Communications, 2019, 40(6): 160-176.
- [3] LI Z M, PENG C, YU G, et al. DetNet: a backbone network for object detection[J]. 2018. arXiv: 1804.06215.
- [4] 李季明, 张宁. 具有随机性的确定性网络模型[J]. 复杂系统与复杂性科学, 2007, 4(2): 56-61.
LI J M, ZHANG N. Deterministic network model with randomness[J]. Complex Systems and Complexity Science, 2007, 4(2): 56-61.
- [5] LUONG N C, HOANG D T, GONG S M, et al. Applications of deep reinforcement learning in communications and networking: a survey[J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3133-3174.
- [6] ROBINSON Y H, JULIE E G, SARAVANAN K, et al. FD-AOMDV: fault-tolerant disjoint ad-hoc on-demand multipath distance vector routing algorithm in mobile ad-hoc networks[J]. Journal of Ambient Intelligence and Humanized Computing, 2019, 10(11): 4455-4472.
- [7] WANG Z, CROWCROFT J. Analysis of shortest-path routing algorithms in a dynamic network environment[J]. ACM SIGCOMM Computer Communication Review, 1992, 22(2): 63-71.
- [8] DAI B, CAO Y Y, WU Z L, et al. Routing optimization meets machine intelligence: a perspective for the future network[J]. Neurocomputing, 2021, 459: 44-58.
- [9] 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 Sur-



veys & Tutorials, 2019, 21(1): 393-430.

- [10] MCKEOWNN, ANDERSONT, BALAKRISHNANH, et al. OpenFlow[J]. ACM SIGCOMM Computer Communication Review, 2008, 38(2): 69-74.
- [11] SCOTT F, HERKE H, DAVID M. Addressing function approximation error in actor-critic methods[C]// Proceedings of the 35th International Conference on Machine Learning. [S.l.: s.n.], 2018:1587-1596.
- [12] UHLIG S, QUOITIN B, LEPROPRE J, et al. Providing public intradomain traffic matrices to the research community[J]. ACM SIGCOMM Computer Communication Review, 2006, 36(1): 83-86.

[作者简介]



伍仲丽（1998-），女，华中科技大学电子信息与通信学院硕士生，主要研究方向为强化学习、路由计算。



曹园园（1996-），女，华中科技大学电子信息与通信学院硕士生，主要研究方向为机器学习、SDN。



黄文睿（1999-），男，华中科技大学电子信息与通信学院硕士生，主要研究方向为图神经网络、网络路由。

戴彬（1977-），男，博士，华中科技大学电子信息与通信学院副教授，主要研究方向为智能路由、边缘计算。

莫益军（1976-），男，博士，华中科技大学计算机科学与技术学院副教授，主要研究方向为智能网络、边缘计算。