

基于P4和机器学习的 路由选择方案探讨

Discussion on Routing Scheme Based on P4 and Machine Learning

李倩,张凯,魏浩然,张娇(北京邮电大学,北京100876)

Li Qian,Zhang Kai,Wei Haoran,Zhang Jiao(Beijing University of Posts and Telecommunications,Beijing 100876,China)

摘要:

传统路由解决方案存在着网络不可感知、收敛慢等缺点。随着数据平面编程语言P4的发展,数据平面具备了实时采集网络状态信息的能力。但大量采集数据通过人为方式难以处理,应用机器学习工具可以很好地解决数据处理困难的问题。在以上2个基础上,通过搭建平台对基于P4和机器学习的路由选择方案进行实验,测试结果表明使用强化学习算法协助路由策略生成的方案表现出良好的性能。由此观之,机器学习在网络领域的潜力巨大,是实现网络智能化的重要工具。

关键词:

P4;强化学习;路由选择;智能化网络
doi:10.12045/j.issn.1007-3043.2018.012.002
中图分类号:TP181
文献标识码:A
文章编号:1007-3043(2018)12-0007-05

Abstract:

Traditional routing solutions have always had shortcomings such as insensitivity to the network and slow convergence. With the development of the data plane programming language P4, the data plane has the ability to collect network status information in real time. However, a large amount of data is difficult to process through human methods, and the application of machine learning tools can solve the problem of difficult data processing. Based on the above two foundations, the P4 and machine learning-based routing schemes are tested by building a platform. The test results show that the scheme using the reinforcement learning algorithm to assist the routing strategy generation has good performance. From this point of view, machine learning has great potential in the field of networking, and is an important tool for realizing network intelligence.

Keywords:

P4; Reinforcement learning; Routing; Intelligent network

引用格式:李倩,张凯,魏浩然,等. 基于P4和机器学习的路由选择方案探讨[J]. 邮电设计技术,2018(12):7-11.

1 概述

近年来,关于是否可以将机器学习方法应用到网络中的讨论吸引了众多网络从业人员的关注。与此同时,许多人也从网络最基础问题——路由问题着手,尝试通过实践检验这一讨论的可行性。

一直以来,路由选择都是网络工程中的重要一环,传统的路由策略一般以最短跳数为转发路径,网络设备没有能力采集数据包在转发过程中经历的具

体信息(如排队时延、队列拥塞状况、流量到达速率等等)^[1]。其次,传统路由策略还存在收敛慢的缺陷,由于相邻交换机之间的信息交换是有一定间隔的,因此当某条链路或节点出现故障时,网络不能及时感知,就会造成大量数据包重传,加剧网络拥塞的情况发生。

软件定义网络的出现,数控分离的思想使得实时获取网络真实状态变为可能。通过基于数据平面可编程语言P4的带内网络遥测框架^[2-3],数据平面可以实时收集数据包在数据平面转发过程中的网络状态信息,并将其封装在数据包头中。这样不仅不需要控

收稿日期:2018-11-09

制平面的干预,还减少了大量网络探测包带来的开销。数据平面收集的大量信息全部交由控制器处理,如此庞大的数据量怎样才能行之有效地转化为路由配置呢?若通过人工建模必将消耗大量的人力和物力资源,而且由于网络状态难以预测,模型建立很难做到普适性。而机器学习擅长解决模型复杂、人工建模困难的问题,因此将机器学习应用到网络工程中不失为一个好的解决方案。

通过搭建实验平台和网络编程,基于P4和机器学习的路由选择方案得到了实验验证。从实验结果来看,经过较短时间的训练,基于P4和机器学习的路由选择方案模型就能迅速收敛,性能表现呈现迅速上升,并超越ECMP的趋势。由此观之,机器学习在网络领域将会发挥巨大的潜能,是促进网络智能化的重要工具。

2 基于带内网络遥测的信息收集

软件定义网络自提出以来,因其使网络拥有者和运营商能够对网络行为进行编程而取得了巨大的成功。然而,其可编程性仅局限于网络控制平面,其转发平面在很大程度上受制于功能固定的包处理硬件。数据平面编程语言P4自2014年提出以来,就受到业界的广泛关注,它完全摆脱了网络数据平面的束缚,让网络拥有者、工程师、架构师及管理员可以自上而下的定义数据包的完整处理流程。

P4语言具有三大特性:协议无关性、目标无关性和现场可重配置能力。协议无关性指的是网络设备不与任何特定的网络协议捆绑在一起。相反,对于任何所需要的网络数据平面协议和数据包处理行为,程序员都可以使用P4进行描述。目标无关性指的是程序员无需关心底层硬件细节,就可以描述数据包的处理逻辑,P4程序经过编译后,会屏蔽掉底层差异。现场可重配置能力指的是程序员在部署交换机后能对数据包的处理方式进行再次修改。

基于P4的带内网络遥测(INT)是一种网络测量框架,旨在允许通过数据平面收集和报告网络状态,而不需要控制平面的干预。在带内网络遥测架构模型中,数据包包含被网络设备解释为“遥测指令”的包头字段。这些指令告诉支持带内网络遥测功能的设备在数据包转发时收集和写入某些状态。信息源(应用程序、终端主机网络栈、虚拟机监控程序、NIC、发送端ToR等)可以将指令嵌入正常的数据包或特殊的探测

包中。信息接收器收集并选择性地报告这些指令的收集结果。这种模式允许监视器在转发时监视数据包所“观察到”的确切数据平面状态。其中,支持带内网络遥测的交换机按照指令写入的交换机自身数据称为元数据。通过对这些元数据的组合和分析,就可以得到很多有用的OAM信息。

带内网络遥测并没有一种规定的封装格式,可以将带内网络遥测头部加在任意封装类型的可选字段或净负荷部分,可用的通用协议栈包括VxLAN、Geneve、NSH、TCP、UDP等。对于不同的封装格式,只需要为其添加一种新的类型号或其他标识,指明后续包头为带内网络遥测包头即可。同样的,使用自定义的封装格式也是完全可行的。

通过带内网络遥测框架,控制器可以收集到数据包在转发过程中所经历的真实网络状态,而不需要发送特殊的测试包用于测试网络状态,从而获得了更精确和可靠的数据。

数据平面收集到的大量数据将会交付给控制平面,经过预处理后,这些数据将作为机器学习的训练集,指导路由策略的生成。

3 强化学习算法PPO

机器学习方法按类型可分为监督学习、无监督学习和强化学习。3种方法各有各的特点和适用范围。

3.1 监督学习

监督学习^[5]是在知道输入和输出的情况下训练出一个模型,将输入映射到输出。在开始训练之前,就已经知道了输入和输出,训练任务就是建立起一个将输入准确映射到输出的模型,当给模型输入新的值时就能预测出对应的输出了。在这一过程中机器不断的通过训练输入来指导算法不断改进。如果输出的结果不正确,那么这个错误结果与期望正确结果之间的误差将作为纠正信号传回到模型,纠正模型。在深度学习中,著名的反向传播算法本质上就是将误差向后传播来指导模型改进的。

3.2 非监督学习

非监督学习与监督学习不同,非监督学习并不需要完整的输入输出数据集,并且系统的输出经常是不确定的。它主要被用于探索数据中隐含的模式和分布。非监督学习具有解读数据并从中寻求解决问题的能力,通过将数据和算法输入到机器中将能发现一些用其他方法无法见到的模式和信息。

3.3 强化学习

强化学习^[6]突破了非监督学习,为机器和软件如何获取最优化的结果提供了一种全新的思路。它将如何最优化主体的表现和如何优化这一能力之间建立起了强有力的链接。通过奖励函数的反馈来帮助机器改进自身的行为和算法。

如图1所示,强化学习中主体通过行为与环境相互作用,而环境通过奖励函数来帮助算法调整做出行为决策的策略函数。从而在不断的循环中得到表现优异的行为策略。它十分适合用于训练控制算法和游戏AI等场景。

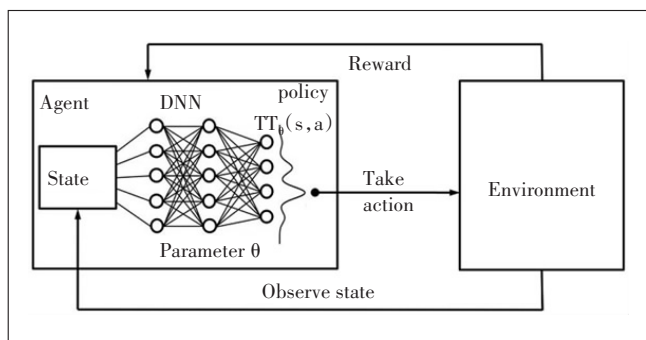


图1 强化学习模型

但强化学习在实践中并不简单,人们利用很多种算法来实现强化学习。简单来说,强化学习需要指导机器做出在当前状态下能获得最好结果的行为。

在网络工程路由策略这一场景中,网络状态瞬息万变,采集信息与路由策略间的关系也十分微妙,难以通过数学建模表达出来,因此,选用强化学习协助路由策略生成是最符合实际场景的方法。强化学习的模型有很多,其中PPO算法^[7]是OpenAI推出的号称效率最高、训练效果最明显、最容易实现的算法,因此实验中选用了PPO算法。

输入。数据平面通过INT框架可采集到的数据有很多种,根据数据与路由选择的相关性分析,实验中选取了交换机队列拥塞程度、交换机排队时延、链路负载3类数据作为PPO算法的输入。

输出。在PPO算法的输出表示上,存在着一些争议。最直接的策略是输出所有的路由表项,这样控制器不需额外处理就可以通过北向接口将路由表项下发到数据平面的P4交换机中,但存在随着网络拓扑的扩展,输出数量将会呈指数增长的问题,不利于网络扩展。另一种策略是输出所有链路的权值,然后控制器将链路权值转化为相应的路由策略。将链路权值

转化为路由策略的处理是相对简单的,在很多论文^[10-11,15]中都论述了相应的方法。同时,随着网络拓扑的扩展,输出规模的增长呈线性,相对于直接输出所有路由表项的方式节省了很多内存空间。实验中选取所有链路的权值作为PPO算法的输出。

反馈函数。强化学习中反馈函数的设计是至关重要的,合适的反馈函数能训练出良好的机器学习模型。反馈函数的设计与优化目标是密不可分的。评价路由算法优劣的性能指标有流完成时间、链路利用率、链路拥塞程度等等。实验中采用链路利用率作为性能指标,优化目标为最小化最大链路利用率。因此,实验中将反馈函数设计为全局链路中最大链路负载与链路带宽的比值的负数,指导机器学习模型的训练。

图2为PPO算法的神经网络结构图。在训练过程中,环境(Environment)为数据平面采集的数据,经过预处理后,采集数据作为actor网络的输入数据,Actor网络根据输入数据(state)输出一组链路权值(action)。Actor网络的输出结果好坏由critic网络评判。根据下一周期的采集数据,critic网络将对Actor网络在前一周期的动作给出评判(reward),指导Actor网络的优化。

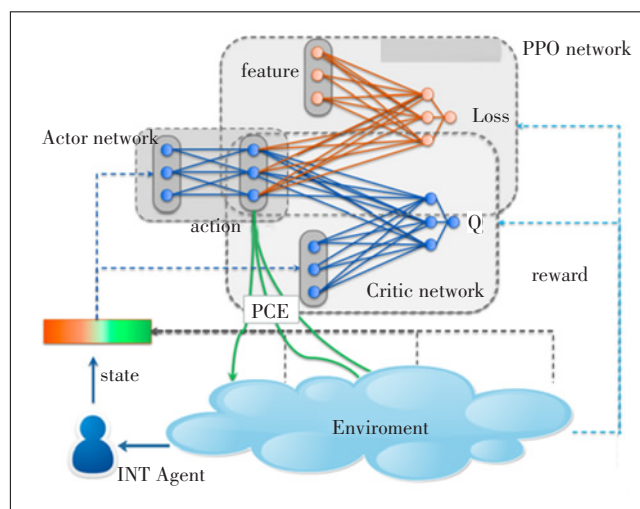


图2 PPO算法神经网络结构图

4 系统框架与性能分析

系统整体框架如图3所示。在数据平面,通过基于P4的带内遥测技术,周期性向数据包添加INT遥测指令,数据包依据遥测指令收集沿途链路和网络设备信息,实时获得网络状态。在INT终端(INT Sink)交换

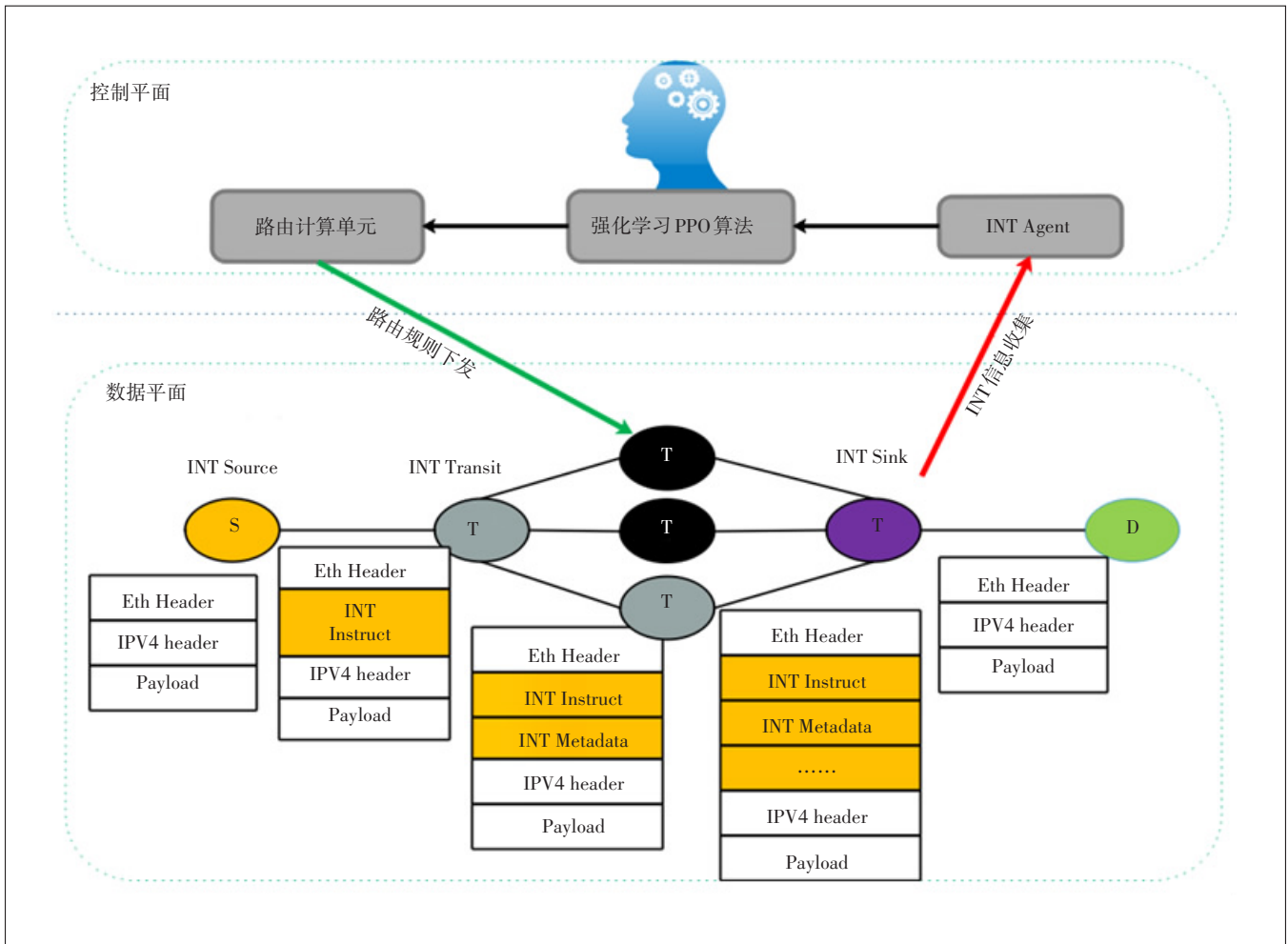


图3 系统架构图

机上,交换机自动将收集到的INT信息从数据包剥离,传给INT监测模块(INT Agent)。在控制平面,INT监测模块(INT Agent)将数据平面获取的当前链路负载、交换机内排队时延、队列拥塞程度等信息经过处理后作为神经网络模型的输入数据,通过PPO算法计算出全局链路权值,继而通过路由计算单元将其转换为路由决策。之后通过P4 runtime将路由决策部署到各个交换机上。

通过实验环境搭建和程序编写,基于P4和机器学习的路由选择方案得到实现。图4为基于P4和机器学习的路由选择方案与ECMP选路方案的性能分析图。红色表示基于P4和机器学习的路由选择方案测试结果,蓝色表示ECMP选路方案的测试结果。横轴表示训练次数,纵轴表示以ECMP方案为基准,两者选路方案测试结果中最大链路利用率的比值。从图4可以看出,经过较短时间的训练,基于P4和机器学习的

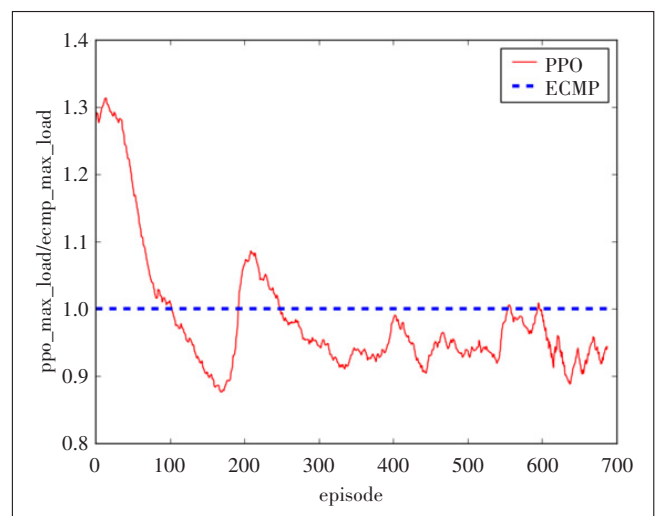


图4 实验方案与ECMP性能分析图

路由选择方案性能就明显优于ECMP,然后之后一段时间,优于机器学习模型不稳定,智能路由方案的性

能出现明显波动,增加训练周期后,模型逐渐稳定,可以看到,基于P4和机器学习的路由选择方案性能大大优于ECMP选路方案。

5 总结与展望

目前,很多领域都尝试着用机器学习的方法作为解决问题的新思路^[4],事实证明,机器学习工具的潜力是非常巨大的。从1997年IBM深蓝击败国际象棋大师卡斯帕罗夫,到2016年谷歌人工智能团队所创造的AlphaGo战胜韩国围棋棋手李世石,人工智能在国际上引起了一次又一次“地震”,吸引了各界人士的关注。

在网络领域,机器学习的“表现”也十分突出。例如拥塞控制^[8-9]、网络瓶颈监测、优化数据中心能源损耗^[13]、资源分配^[14]、视频流的比特率选择^[15]等。其中,在路由选择方面,出现了用强化学习Q-learning算法指导路由的Q-routing方案。Q-routing方案是基于分布式的,每个路由器需要单独学习从数据包头部到出端口的映射关系,并且还要不断交换解析的数据包中对应不同目的地址的时延信息。Q-routing方案是一种去中心化的、数据包粒度的路由选择方案,在可扩展性和通信开销方面有很大的挑战。除此之外,在2017年HotNets会议上发表的《learning to route》^[1]论文也对将机器学习工具应用到路由选择中进行了详细的探讨。文中比较了监督学习和增强学习2种方案,实验结果表明基于增强学习TRPO算法的softmin routing路由方案表现出较好的效果,但相较于不添加机器学习的路由方案,基于机器学习的方案性能优势没有表现的特别明显,随着训练次数的增多,性能优势可能会逐渐增大。

网络因其状态瞬息万变而难以预测,传统的应对措施只能通过人为建模建立确定性关系来解决问题。例如,在路由选择方面以最短跳数作为路由表项更新的依据;在资源调度方面,通过计算最优静态阈值的方式作为缩扩容的依据。实际上,网络中各个因素之前的相关性是不能简单地通过数学表达式的形式表达出来的,至少目前是如此。通过将机器学习工具应用到网络领域,以反馈的方式训练模型,以无限逼近的方式寻求最优结果,显然是一个更合理的方法。基于P4和机器学习的路由选择方案是将机器学习工具应用到网络中的一个初步实践,可优化和可扩展空间还有很多,相信在不久的将来,机器学习在网络领域

的成果会越来越多,人工智能将引领一个智能化时代的诞生。

参考文献:

- [1] VALADARSKY A, SCHAPIRA M, SHAHAF D, et al. Learning to route[EB/OL].[2018-08-20]. https://www.researchgate.net/publication/321325749_Learning_to_Route.
- [2] The P4 Language Consortium. The P4 Language Specification[S/OL].[2018-08-20]. <https://p4.org/>.
- [3] The P4.org Applications Working Group. In-band Network Telemetry (INT) Dataplane Specification[S/OL].[2018-08-20]. <https://p4.org/>.
- [4] MAJER M, BOBDA C, AHMADINIA A, et al. Packet Routing in Dynamically Changing Networks on Chip[C]// IEEE International Parallel & Distributed Processing Symposium. IEEE, 2005.
- [5] MOHRI M, ROSTAMIZADEH A, TALWALKAR A. Foundations of Machine Learning[M]. The MIT Press, 2012.
- [6] SUTTON R S, BARTO A G. Reinforcement learning: An introduction[J]. IEEE Transactions on Neural Networks, 1998, 9(5): 1054.
- [7] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal Policy Optimization Algorithms[J]. OpenAI, 2017.
- [8] DONG M, LI Q, ZARCHY D, et al. PCC: Re-architecting Congestion Control for Consistent High Performance[J]. NSDI, 2015.
- [9] WINSTEIN K, BALAKRISHNAN H. Tep ex machina: Computer-generated congestion control[J]. SIGCOMM, 2013.
- [10] FORTZ B, THORUP M. Increasing Internet Capacity Using Local Search[J]. Computational Optimization & Applications, 2004, 29(1): 13-48.
- [11] MICHAEL N, TANG A. HALO: Hop-by-Hop Adaptive Link-State Optimal Routing[J]. IEEE/ACM Transactions on Networking, 2014, 23(6): 1-1.
- [12] XU D, CHIANG M, REXFORD J. Link-State Routing with Hop-by-Hop Forwarding Can Achieve Optimal Traffic Engineering[C]// INFOCOM 2008. The 27th Conference on Computer Communications. IEEE. IEEE, 2008.
- [13] DeepMind AI Reduces Google Data Centre Cooling Bill by 40%[EB/OL].[2018-08-20]. <https://goo.gl/QTdU2T>.
- [14] MAO H, ALIZADEH M, MENACHE I, et al. Resource management with deep reinforcement learning[J]. HotNets, 2016.
- [15] MAO H, NETRAVALI R, ALIZADEH M. Neural adaptive bitrate streaming with pensive[J]. SIGCOMM, 2017: 197-210.

作者简介:

李倩,北京邮电大学在读硕士,主要研究方向为数据平面编程语言P4和软件定义网络;张凯,北京邮电大学在读硕士,主要研究方向为机器学习以及网络人工智能;魏浩然,北京邮电大学在读硕士,主要研究方向为数据中心以及网络人工智能;张娇,毕业于清华大学,副教授,博士,主要研究方向包括数据中心网络、网络功能虚拟化、网络人工智能、未来网络体系架构等。