Network Planning with Deep Reinforcement Learning

雷雨诗-19211330-阅读笔记M3

目录

[一、文献信息 1](#_Toc103890510)

[二、问题意义 1](#_Toc103890511)

[2.1研究问题 1](#_Toc103890512)

[2.2研究背景 2](#_Toc103890513)

[2.3研究意义 2](#_Toc103890514)

[三、思路方法 2](#_Toc103890515)

[3.1网络规划的基础知识 2](#_Toc103890516)

[3.2网络规划中需要重点考虑的问题 2](#_Toc103890517)

[3.3神经计划 3](#_Toc103890518)

[3.3.1 神经计划对问题的解决： 3](#_Toc103890519)

[3.3.2 神经计划的工作流程： 3](#_Toc103890520)

[3.3.3 神经计划训练算法原理： 4](#_Toc103890521)

[3.3.4 神经计划的优化： 4](#_Toc103890522)

[四、实验结论 4](#_Toc103890523)

[4.1神经计划的优化效率如何？ 4](#_Toc103890524)

[4.2 神经计划为小规模的问题生成的网络规划的最优性 5](#_Toc103890525)

[4.3 对于大规模问题神经计划的扩展性 5](#_Toc103890526)

[4.4 神经计划中其它参数如何影响结果 6](#_Toc103890527)

[五、启发思考 6](#_Toc103890528)

一、文献信息

作者：[Hang Zhu](https://xueshu.baidu.com/s?wd=author%3A%28Hang%20Zhu%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[Varun Gupta](https://xueshu.baidu.com/s?wd=author%3A%28Varun%20Gupta%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[Satyajeet Singh Ahuja](https://xueshu.baidu.com/s?wd=author%3A%28Satyajeet%20Singh%20Ahuja%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[Yuandong Tian](https://xueshu.baidu.com/s?wd=author%3A%28Yuandong%20Tian%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[Ying Zhang](https://xueshu.baidu.com/s?wd=author%3A%28Ying%20Zhang%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)，[Xin Jin](https://xueshu.baidu.com/s?wd=author%3A%28Xin%20Jin%29%20&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dperson)

论文题目：Network Planning with Deep Reinforcement Learning

发表途径：SIGCOMM '21: ACM SIGCOMM 2021 Conference

发表时间：2021/08/09

二、问题意义

## 2.1研究问题（论文摘要）

本文主要研究的是一种**神经计划（NeuroPlan）**。这是一种用深度强化学习（RL）的方法来解决网络规划的问题。我们在研究这个NeuroPlan的过程中会涉及到多步骤的决策制定问题和成本最小化问题，所以我们很自然的会想到用深度强化学习的方法来解决这个问题。首先，我们使用图神经网络(GNN)和一种新的基于特定领域的节点链路转换（将节点拓扑转换为链路拓扑）进行状态编码，以处理规划决策过程中演化的网络拓扑的动态特性。然后，我们利用一种两阶段混合方法，首先使用深度RL来修剪搜索空间，然后使用整数线性规划（ILP）求解器来找到最优解。

## 2.2研究背景（通过论文introduction总结）

**a.网络规划的重要性。**

在今天的社会下，互联网飞速发展，人们的生产生活已经离不开网络，所以我们需要有大规模的网络服务来支持今天的社会发展。而大规模的网络服务依赖于高性能、可靠的广域网。近年来新兴的技术如人工智能、海量物联网等等，都对广域网的宽带需求持续快速增长。所以说，在这样的情况之下，网络规划是规划和升级广域网的一个常规的、关键的过程，以满足互联网服务的性能和可靠性需求，同时将成本降到最低。

**b.神经计划——基于深度强化学习来解决网络规划问题。**

网络规划问题很适合用深度强化学习来解决。其主要原因有两点：第一，网络规划需要在IP层和光学层做出多个决策，这种多步骤决策的问题正是强化学习被设计用来解决的一类问题。第二，多步骤决策的每一步影响是隐性的，只能在整个网络规划完成后进行评估，而深度强化学习可以被设计为一种延迟奖励的形式，将影响变成显性，去显性地优化全局目标。

## 2.3研究意义（通过论文introduction总结）

论文研究的神经计划使用延迟奖励等机制，可以避免人们在对网络最优性和复杂度之间进行权衡时耗费的大量人力物力，并且也很容易与其他方法相结合。此外，论文研究的神经计划通过在大型生产网络和真实拓扑中的评估中发现，其相比于人工调节的方法还可以降低高达17%的网络成本。最后，在如今互联网飞速发展的趋势下，本文提出的神经计划可以在人工智能、自动驾驶网络等新兴的领域有很好的应用，为我们实现目标迈出了具体的一步。

三、思路方法

本文研究神经计划，首先给出了网络规划层面中一些基础性的概念知识，然后从其关键技术——深度强化学习的局限性入手，进而给出我们的解决方法，最后通过实验对我们的研究进行评估，以得出最后成果，具体如下：

## 3.1网络规划的基础知识

网络规划的最终目标是确定如何改变网络规模以满足给定未来需求预测的流量。它有如下性质：

首先，**网络规划具有跨层性质**，它同时考虑了IP层和光学层，跨层布局IP链路和底层光纤路径的映射。具体是：在IP层，网络规划决定了每个IP链路的容量以及转发器和路由器的数量，以在每个节点取得以支持这样的容量。在光学层，网络规划需要估计用到的光纤数量以及构建的新光纤的路径。

其次，一次性做出所有网络规划的决定是很复杂的，而且不能满足业务的需要，所以网络规划需要是一个**多阶段的迭代过程**。在网络的生成中，规划过程需要分别在短期和长期范围内进行，来为运营团队产生短期的可操作性信号，同时让网络在长期上往最优网络的方向上发展。短期规划提供了在未来短期的时间范围内，在给定光纤、现有IP链路上增加或移除链路容量的决策。长期规划决定了未来的光纤路径、硬件设备和新的节点。短期与长期规划的目标都是提取最具成本效益的拓扑。

最后，网络规划是具有故障感知能力的，它的一个关键的目标就是确保骨干网络在不同的故障场景下网络的**可靠性**。具体还是通过短期规划（决定现有IP链路容量）与长期规划（添加或改变路径）来实现。

## 3.2网络规划中需要重点考虑的问题

网络规划问题可以被描述为一个线性规划问题，具体是在流量守恒、链路容量、频谱效率等约束条件下求解网络成本的最小化方法。在求解线性规划问题是可以直接用一个现成的ILP求解器，但ILP求解器的**扩展性**是目前面临的主要问题。当拓扑中的节点链路以及故障场景增加时，这个问题可以被转换成一个有上万个变量以及约束条件的ILP问题，这样的计算无论是从时间还是成本上来说都是具有很大挑战的，而从目前的现状来看，拓扑结构的大小还在以每年20%的速度增长。

由于所需的计算量呈指数增长，因此ILP方法不能直接用于大型拓扑。在这种情况下，我们依靠**人工协调的探索式方法**来解决问题，具体如拓扑分解、故障选择等。拓扑分解是将拓扑分解为几个较小的子拓扑，再将每个子拓扑都用ILP求解。故障选择是指选择全部故障场景的一个子集以减少线性规划求解释的约束和变量。

但在应用人工协调的探索式方法时也存在问题：虽然它们的目标都是修剪探索空间，使问题易于处理，但它们都依赖于人的专业知识与操作经验，这是一个需要时间与手动试错的过程。同时，当搜索空间被修剪时，在解的最优性和ILP的可处理性之间存在一个基本的权衡，这个权衡又另需要通过迭代过程去寻找到，无疑又是**增加了问题的复杂性**。此外，这种方法还具有局部最优性，如何通过局部最优寻找到全局最优也是一个问题。

## 3.3神经计划

神经计划所基于的深度强化学习方法，可以自然地映射到网络规划问题上：其中环境是具有流量需求和可靠性策略的网络拓扑，动作是改变拓扑的链路容量，奖励是网络拓扑的最终成本，以及在可靠性策略下是否满足流量需求。

此外，在探索式方法的局部最优性质上，深度强化学习能够选择一个影响只能在以后评估的行动，直接优化全局的目标。

### 3.3.1 神经计划对问题的解决：

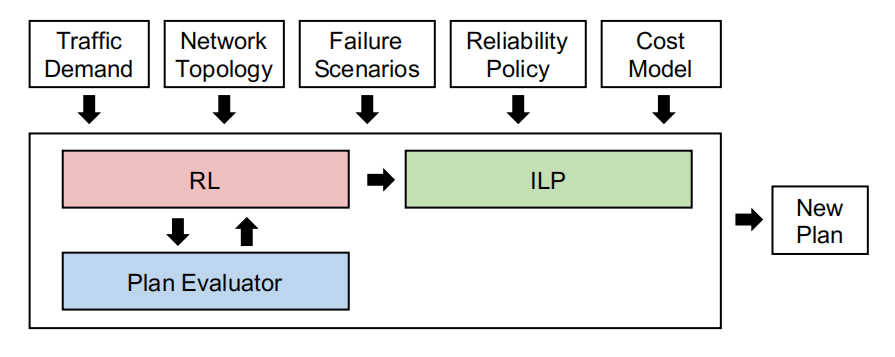
a.虽然将深度RL应用于网络规划是很自然和有吸引力的，但它的一个挑战是如何在深度强化学习中进行**动态网络拓扑的建模**：深度强化学习通常需要一个向量来将环境状态编码为其神经网络的输入，以生成动作，而网络规划问题的环境是网络拓扑结构图，不容易用向量表示。

我们用GNN来解决这个问题。用GNN对网络拓扑进行编码，它可以以一个图作为输入，并通过学习生成嵌入向量来表示其节点，链路或整个图表，这也就解决了动态的表示问题。我们以动态网络拓扑作为GNN的输入，生成整个拓扑的嵌入向量，捕获拓扑的基本特征，并作为深度强化学习的输入。

b. **最优性和可处理性之间的权衡。**在此问题上我们利用了RL和ILP的两阶段混合方法。在第一阶段，我们使用深度强化学习来寻找一个合理的解决方案。在第二阶段，我们使用ILP搜索深度强化学习方法找到的解附近的空间。在第二阶段我们同时使用松弛因子α来控制ILP探索空间的大小。

### 3.3.2 神经计划的工作流程：

下图显示的是神经系统计划的工作流程。他有五个输入，分别是流量需求、网络拓扑结构、故障场景、可靠性策略和成本模型。图中深度强化学习算法（RL）只需要对网络拓扑进行编码，其他四个组成部分则由计划评估器来处理。RL与计划评估器进行交互，以学习生成能够最小化网络成本，同时满足可靠性策略下的流量需求的网络计划。计划评估器向RL产生“奖励”以不断优化网络。“奖励”是根据需求是否得到满足和计划的成本的组合来计算的。在学习过程完成后，RL输出第一阶段的初始计划。在第二阶段，ILP求解器使用由松弛因子α定义的初始解附近的修剪搜索空间中找到最终解。



### 3.3.3 神经计划训练算法原理：

算法首先初始化参与节点、GNN、和epoch的相关参数。对于每一个epoch，需要从当前的状态中返回的动作概率的分布中抽样出几个网络规划。对于同一状态可以抽样出多个不同的规划以实现充分的搜索。

每一个规划的生成我们都把它叫做是一个轨迹。NeuroPlan从具有原始链路容量的网络拓扑结构开始，对于每一个轨迹，通过迭代估计下一步动作，生成网络规划，直至轨迹结束。

轨迹结束的条件：1，当前网络拓扑满足可靠性策略下的流量需求。2，轨迹长度超过预定义的阈值。3.轨迹被当前epoch切断。满足条件之一即可停止。

### 3.3.4 神经计划的优化：

**a.源聚合。**我们在构造线性规划问题的约束条件时，不是使用单个流来构造约束，而是使用源聚合。源聚合主要是将具有相同源的流聚合为单个流，即将相似行为归为一类，这种优化减少了约束的数量，从而减少了线性规划问题的复杂度。

**b.有状态的故障检查。**我们知道，如果一个网络在一次故障中存活下来，那么一个容量更大的网络就可以保证在同样的故障中存活下来。我们的方法是，保持了一个固定的故障顺序，对于每个步骤，只检查从上一个步骤中没有存活下来的故障，而不是试图一次解决所有故障。同时，我们还可以对故障进行分组，并使用多台机器并行检查故障组，从而可以对有大量故障的问题进行训练。进行有状态的故障检查后，我们只更新模型中受故障影响的约束，避免为每个故障从头开始构建模型，降低复杂度。

**c.其他。**我们利用AI加速器(如GPU)加速训练的能力，同时采用更大的容量增量单元，使深度强化学习算法能够以更少的步骤得到可行的解决方案，并将数据放入GPU内存中。

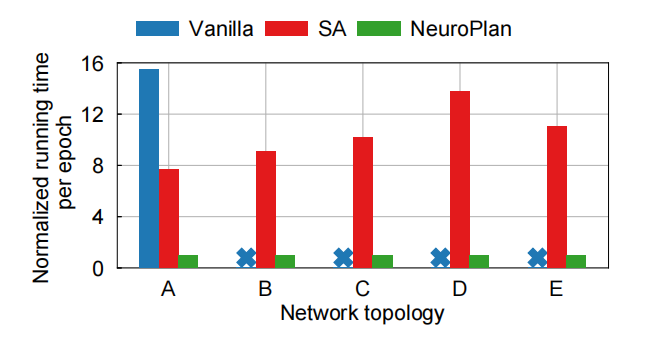
四、实验结论

我们使用了五种不同尺度的生产网络拓扑，即A、B、C、D和E，按拓扑大小的升序列出。在规模上，A有数十个IP链路，数十个故障和数十个流，需要增加一些Tbps容量才能实现可行的解决方案；E有数百个IP链路，数百个故障和大约1000个流，需要增加几百Tbps容量才能实现可行的解决方案。我们使用来自生产网络的真实故障场景、流量需求、可靠性策略和成本模型。同时我们还报告了由深度强化学习产生的第一阶段的结果。

实验主要围绕以下几个问题：

## 4.1神经计划的优化效率如何？

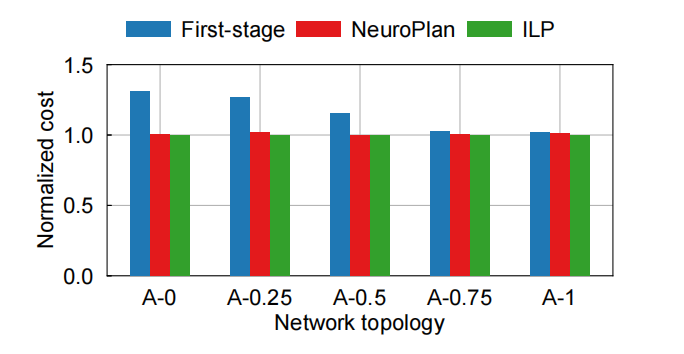
根据上文我们提到神经计划的主要优化有源聚合以及状态故障检查。我们比较了普通计划评估器（Vanilla）、具有源聚合的计划评估器（SA）、同时具有源聚合和状态故障检查的计划评估器（NeuroPlan），分别在五种拓扑下（A、B、C、D、E），10个epoch的运行平均运行时间。我们可以看到如下结果：



对于拓扑A，SA直接将运行时间缩短了近两倍，而对于其他拓扑，SA的效率更高（因其显著减少了线性规划的约束条件的数量）。对于NeuroPlan，它更比SA快了7到14倍，因为它还有状态故障检查优化。

## 4.2 神经计划为小规模的问题生成的网络规划的最优性

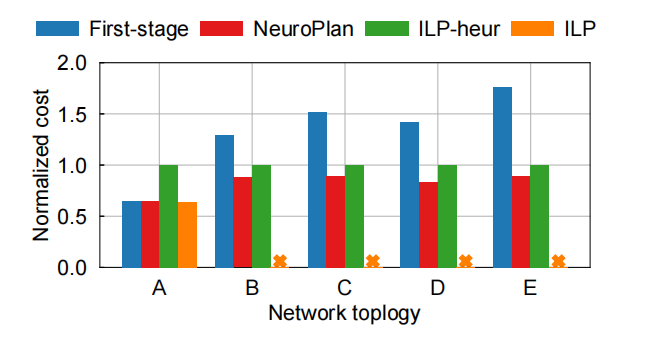
我们知道ILP可以为小规模问题生成最优解，所以比较了神经计划和ILP在小规模问题上的作用，以评估神经计划的最优性。我们改变小拓扑A的原始容量，以创建不同搜索空间大小的多个合成问题。具体来说，A-0、A-0.25、A-0.5、A-0.75和A-1表示每个链路的原始容量分别为拓扑A上对应链路的原始容量的0%、25%、50%、75%和100%。我们将松弛因子α设为2。结果如下所示：



原始容量分别为75%(A-0.75)和100%(A-1)的拓扑，其第一阶段（First-stage）的结果已经接近最优（蓝色柱状图与红色绿色柱状图高度相差不大）。即使使用0%(A-0)的原始容量，其中深度强化学习算法从零开始生成网络规划，第一阶段的成本结果也不超过最优成本的30%。在第二阶段之后，神经计划能够以不超过最佳成本的2%的成本产生最终的计划。由此看出NeuroPlan的最优性。

## 4.3 对于大规模问题神经计划的扩展性

NeuroPlan可以克服ILP所遇到的可扩展性问题。所以比较了第一阶段、神经计划、ILP-heur（探索启发式ILP）、ILP在五个拓扑下的结果。

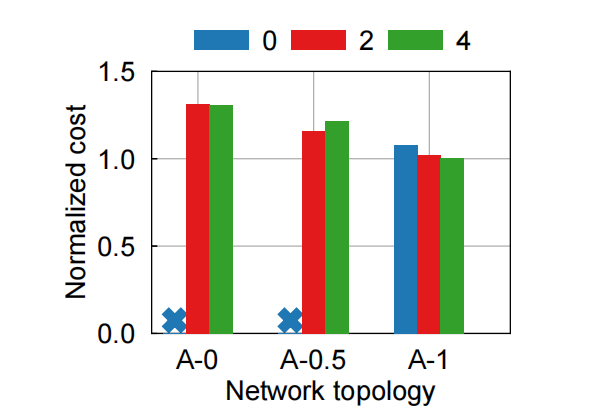


ILP只能解决拓扑A的问题，因为其他拓扑的规模比A大得多，而ILP不能在合理的时间内解决。ILP-heur未能在所有拓扑的最优性和可处理性之间实现良好的权衡，因为不同的问题需要不同的权衡才能获得良好的结果，而不同拓扑ILP-heur的归一化成本是一样的。相比之下，NeuroPlan会针对每个问题自动修剪搜索空间。在得到修剪后的搜索空间后，NeuroPlan可以在第二阶段轻松地得到最优解。此外，与ILP-heur相比，NeuroPlan可以以11%的成本获得大型网络拓扑(即B、C、D和E)的计划，成本从17%降低到11%。

## 4.4 神经计划中其它参数如何影响结果

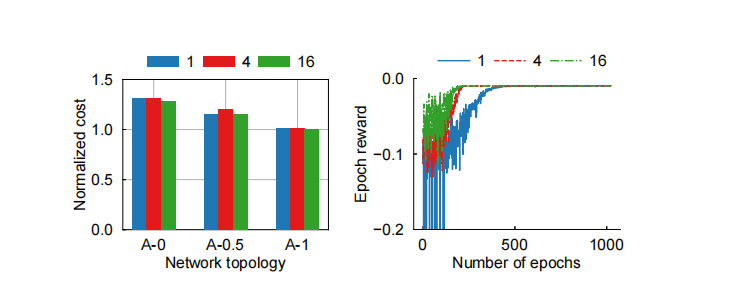
我们对三种合成拓扑，即A-0、A-0.5、A-1进行了不同参数的敏感性分析，改变的参数如GNN层数、每步最大容量单位、松弛因子α等。

1. **GNN层数不同的实验结果：**



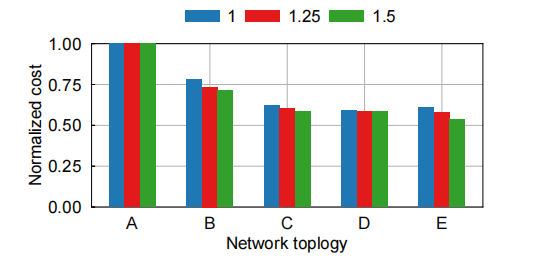
上图显示了第一阶段归一化的成本及其相应的最优成本。在A-1拓扑下，即使没有GNN，神经计划也可以学习，然而，它不能收敛于其他的拓扑结构(即A-0和A-0.5)。这表明深度强化学习的多层感知器可以处理简单的问题，而GNN对于更大规模的问题至关重要。两层或四层的GNN也有类似的结果。

**b.每步最大容量单位的影响实验结果：**



上左图显示每步的最大容量单位对第一步实验的结果几乎没有影响。上右图显示了A-1拓扑时，不同epoch数量实验的收敛速度。我们发现较大的最大容量单位导致相对于epoch数的收敛速度。实验还尝试了其他拓扑，并发现这个结论不适用于其他拓扑。最终的结果是：较大的最大容量单元只有利于容量增量集中在几个链路上。

**c.影响因子α不同的实验结果：**



我们使用不同的α(1、1.25、1.5)评估五种拓扑。通过上图可以发现，有一个更大的α，神经计划可以在更大的搜索空间中得到更好的解决方案。因此，影响因子为最优性和可处理性之间的权衡提供了一个可调的旋钮。

五、启发思考

通过这篇论文，除了学习NeuroPlan以及其实验结果，其实我们对网络规划的一些基本情况也有了较为具体的了解。我们了解了网络规划的最终目标，即成本优化——在满足服务期望的同时，最小化网络成本。同时，容量规划对网络的性能、可靠性和成本也也至关重要。对于如何解决次问题，论文很具体地讲述了如何将网络规划问题转换为一个线性规划问题，这对于我们将现实情况中的问题与数学算法结合起来解决问题有了很大的启发，其实在显示生活解决问题的过程中，我们需要的正是这种问题的归纳与转化能力。本文提出的除了是一种针对网络规划问题具体的神经计划，其实他也给出相关问题的技术层面的解法，即并利用GNN和深度强化学习算法将网络成本最小化，其实这也是在考研我们对问题的举一反三能力。

关于GNN与深度强化学习算法的一些思考。在本文中NeuroPlan利用GNN的能力对网络拓扑进行编码，并利用强化学习有效地探索状态空间，修剪线性规划问题的搜索空间。但其实这两者本生就有很广泛的应用。GNN是一种高效的图表示方法，可以广泛应用于文本分类、分子特征提取、蛋白质结构预测、芯片设计和节点分类等多个领域。对于强化学习，在现在很多解决问题的方法已经将强化学习应用于优化问题上，包括学习切割整数规划，学习混合整数线性规划的分支策略，构建ML编译器和优化神经网络架等等。将GNN和强化学习结合起来，还可以应用于更多不同的任务，如解决经典的组合问题和生成分子图等等。我们可以看出，其实这些方法都是近几年的热点，新技术的出现必带动科技的飞速发展。

在未来，我们希望神经计划能够激发更多将人工智能技术应用于网络自动化中的应用，也同样希望人工智能领域能够进一步的发展，为人类社会带来便利。科技在发展，时代在进步，我们需要不断探寻与发现，伫立时代的全新浪潮。