**Reliability-aware Dynamic Service Chain Scheduling in 5G Networks based on Reinforcement Learning**

摘要--作为未来5G网络的关键推动因素，服务功能链（SFC）沿着虚拟网络功能（VNF）的链条转发流量，以提供网络服务的灵活性。SFC中最重要的问题之一是在计算节点之间部署VNFs和安排到达的请求，以实现低延迟和高可靠性。现有的工作考虑了一个静态的网络，并假设所有的SFC请求都是事先知道的，这是不现实的。在本文中，我们专注于动态5G网络环境，其中SFC请求随机到达，计算节点可以重新部署所有类型的VNF，并有时间成本。我们将支持NFV的5G网络中的SFC调度问题表述为混合整数非线性编程。其目标是使满足延迟和可靠性约束的请求数量最大化。为了解决这个问题，我们提出了一种有效的算法，以决定VNF的冗余，同时使延迟最小化。然后，我们提出了一种最先进的强化学习（RL）来学习SFC调度策略，以提高SFC请求的成功率。我们的方法的有效性通过广泛的模拟进行了评估。结果显示，我们提出的RL解决方案可以比基准提高18.7%的成功率。

I. 前言

最近，软件定义网络（SDN）和网络功能虚拟化（NFV）在网络架构演进中发挥了重要作用。传统上，网络功能是由专用的硬件设备（中间盒）实现的，这增加了网络服务提供商的资金和运营成本，同时也造成了网络耦合的问题。当一种新的服务出现时，相关的硬件设备必须按照一定的顺序进行部署和连接。这种手工操作极其耗时、耗力和容易出错，阻碍了服务添加和网络升级的操作。为了应对这一挑战，NFV将网络功能的实现从专用硬件转移到基于软件的组件，命名为虚拟网络功能（VNF）[1]。VNF是抽象的网络功能，在一般商品（如基于X86的系统）服务器上运行，没有特定的硬件设备。

在NFV框架中，几个VNF实例的有序组合包括一个服务功能链（SFC）[2]。通过不同VNF序列的流量将建立多个SFC，可以支持各种网络服务，如防火墙、负载均衡器、深度包检测（DPI）、入侵检测系统（IDS）等。NFV实现了SFC中使用的软件实现的网络功能的虚拟化。NFV被SFC采用，以提供高效和有效的网络功能的部署和协调[3]。IETF SFC工作组（RFC 7665）和开放网络基金会（ONF）提出了SFC架构规范，展示了运营商网络、移动网络和数据中心网络中的用例。

许多研究对SFC调度问题进行了广泛讨论，但由于5G网络的低延迟和高可靠性要求，执行SFC调度具有挑战性。为了提高网络服务的可靠性，一个SFC的执行需要有额外的冗余VNF实例。冗余实例会占用更多的计算资源，因此会导致其他网络服务的等待时间增加。

其他网络服务的等待时间。因此，5G环境下的SFC调度器必须更加智能，以平衡延迟和可靠性。在本文中，我们考虑了在动态5G网络环境中的在线调度，其中SFC请求随机到达。在这种情况下，计算节点应该切换其部署的VNF类型，以满足传入的SFC请求中的新类型VNF。然而，在计算节点上重新部署各种VNF并不是无成本的；节点必须关闭之前的VNF并设置新的VNF，这就引入了额外的重新部署时间。以前提出的网络模型假设一个节点在所有的时间段内只能承载单一类型的VNF。他们没有考虑计算节点上VNF的重新部署。这种静态部署策略导致在动态变化的SFC请求下，网络资源的利用效率低下。此外，一些现有的研究考虑了计算节点上的重新部署，但忽略了重新部署的时间成本。如果重新部署经常发生，就会在重新部署上花费大量的时间，增加请求的延迟。因此，解决方案应该平衡延迟和重新部署的问题。为了解决这个问题，我们提出了一个有效的方法，将冗余确定与SFC调度分开。对于前者，我们提出了一种启发式算法来决定VNF的冗余，同时使延迟最小化。对于后者，我们开发了一种基于策略梯度的强化学习（RL）方法，将这些VNF放在计算节点上，目的是使SFC请求的成功率最大化。此外，我们设计了一个新的RL动作，不同于现有的将强化学习应用于调度问题的研究。在相关工作中，行动决定了哪个节点应该部署VNF，其中行动空间的大小与计算节点的数量是固定的。我们方法中的RL行动只是决定是否推迟VNF的执行。因此，我们的RL模型可以扩展到任意数量的计算节点，而无需修改RL网络模型。仿真结果显示，与基准算法相比，我们提出的方法可以提高18.7%的成功率。本文的主要贡献总结如下。

- 据作者所知，本文是第一个考虑在SFC请求随机到达的动态网络环境中具有延迟约束的可靠性感知的SFC调度问题的工作。这种环境中的计算节点可以承载所有类型的VNF，当切换所承载的VNF类型时，需要花费重新部署的时间。

- 本文将可靠性感知的调度问题表述为混合整数非线性编程（MINLP）问题，该问题为NP-hard，显示了SFC调度问题的复杂性和找到全局最优解的难度。

- 我们提出了一种强化学习算法来提供一个可行的调度。我们的方法的有效性是通过综合模拟来揭示的。结果表明，所提出的方法在SFC成功率方面优于其他基准。

算法：

我们提出了一种最佳冗余确定算法，如算法1所示。考虑到VNF的最大冗余数Rm，SFC的长度L，服务功能链S，计算节点的可靠性θ，以及SFC请求的可靠性要求Θ，该算法输出一个冗余列表A，表示SFC中每个VNF的冗余数，其中Redundancy(si)是si的VNF实例数，Rel(A)计算A的可靠性，记得我们认为所有计算节点在其执行中具有相同的可靠性θ。例如，如果有一个SFC，其中的VNF和所需的可靠性Θ如图4所示。一个计算节点的可靠性θ被设置为0.96。冗余列表A被初始化为1，然后从左到右循环递增，直到满足所需的可靠性Θ。之后，我们按计算负荷从高到低对VNF进行排序。在这种情况下，冗余VNF的数量与冗余列表中的数量相对应。算法1确保冗余列表A中两个数字之间的差值小于1，这可以证明是给定冗余量时的最大可靠性。

C. 基于规则的节点选择方法

调度规则被用来为VNF选择计算节点。现有的研究提出了一些精心设计的基于规则的算法。Zhang等人[22]提出了一种优先级驱动的加权算法来解决VNF放置问题，该算法可以低成本地找到一个接近最优的解决方案。在[23]中，作者开发了一种两阶段的算法来最小化VNF部署的总成本。然而，现有的SFC调度解决方案并不适用于本文的问题，因为它们考虑的是静态网络，而且所有的SFC请求都是事先知道的。在这种情况下，我们考虑基于规则的FJSP方法，如最早完成（EFF）和最早开始（ESF）。EFF选择能够提前完成VNF的节点。ESF则选择较早开始使用VNF的节点。这些简单的基于规则的方法只提供了可行的解决方案，我们将在后面提出强化学习的方法来提高其性能。

D. 强化学习用于动态SFC调度

我们开发了一个强化学习（RL），使用一个被称为策略网络的神经网络来动态优化SFC调度。强化学习的代理输入证监会的当前状态并输出一个调度行动。我们的RL框架如图5所示。在基于NFV的5G网络环境下，调度代理观察SFC请求和计算节点状态以决定调度行动。该代理在执行行动后获得奖励，并更新网络参数。这里，代理使用嵌入网络将原始状态信息向量编码为RL网络的低维特征。

调度事件。我们的调度模型是任务触发的。未调度的VNF在一个队列中等待。RL代理 图5. 当VNF等待队列不是空的时候，应该做出一个调度决定。该队列的行为如下。

- 队列的优先级是基于距离最后期限的剩余时间。剩余时间较少的SFC的VNF在队列中具有较高的优先级。

- 新到达的SFC的第一个VNF（连同其所有冗余）将被添加到队列中。

- 当一个VNF（及其所有冗余）完成后，其后续任务将被添加到队列中，除非前者是最后一个任务。此外，如果一个VNF的所有冗余失效，我们认为这个VNF的SFC被中止了

而后续的VNF将不会被添加到队列中。

- 如果其中一个VNF冗余完成，其他的将从队列中移除，如果它们已经被部署在节点上，它们的执行将被中止。

- 位于队列头部的VNF将被调度员（或RL代理）首先调度。