一种结合节点重要度的单人（Single-Player）Monte Carlo树搜索方法first

# 摘要(ABSTRACT)

作为网络虚拟化的关键技术之一，虚拟网络映射（VNE）研究如何有效地将物理资源分配给虚拟网络请求。由于VNE问题的NP难问题特征，大多数现有的方法是启发式算法，往往收敛到局部最优解，并且性能较低。本文提出了一种结合基本蒙特卡罗树搜索（MCTS）方法和节点重要性的VNE-SPMCTS算法来应用领域特定知识。对于虚拟网络请求，我们首先将映射过程建模为有限马尔可夫决策过程（MDP），其中每个虚拟节点按照我们设计的节点重要性顺序映射到一个状态中。然后应用最短路径算法将链路映射到终端状态，并将代价作为奖励的一部分返回。由于MDP的奖励延迟机制，链路映射（link mapping）的结果会影响前一个节点映射阶段的动作选择，使两个映射阶段协调。利用节点重要性，可以将特定领域知识用于MCTS的扩展和仿真阶段，加快搜索速度，更准确地估计仿真值。与现有的经典算法相比，实验结果表明，本文提出的算法能够提高VNE的物理节点平均利用率、接受率、长期收益与成本比的性能。

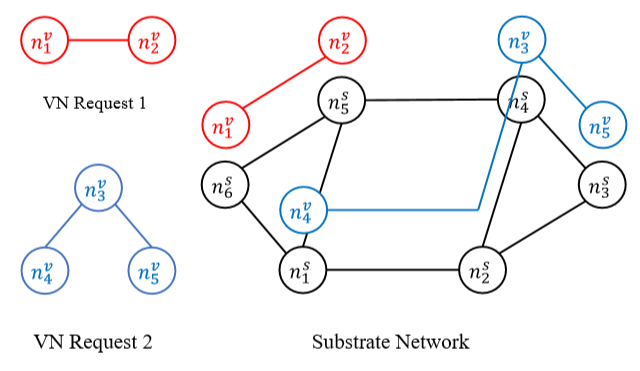
索引词(INDEX TERMS)：网络虚拟化，虚拟网络映射，强化学习，马尔可夫决策过程，蒙特卡洛树搜索，节点排序

# 1.介绍(INTRODUCTION)

过去的30年，网络在支持各种网络技术和分布式应用方面取得了难以置信的成功。然而，由于其多提供商性质(multi-provider nature)，对Internet体系结构的改变已经变得越来越困难，因此，网络虚拟化已经提出了允许多个异构网络体系结构共存于共享底层物理网络[1]。

在网络虚拟化中，传统网络服务提供商(traditional Internet Service Providers (ISPs))的角色分为两个独立的部分：基础设施提供商(Infrastructure Providers(InPs))，管理物理基础设施；服务提供商(Service Providers (SPs))使用多个基础设施提供商提供虚拟网络服务[2]。

在多虚拟网络请求中，物理资源的有效分配和调度(scheduling)对于使得共存的虚拟网络的数量最大化、增加基础设施提供商的利用率和收益是重要的[3]。如图1所示，虚拟网络映射问题尝试以一种消耗最少资源的方式，将虚拟网络的节点和链路映射到底层物理网络上。



虚拟网络映射

使用确定的解决方法，比如整数线性规划(ILP)和混合整数线性规划(MIP)来解决虚拟网络映射[4][5]，然而解决混合整数线性规划问题是NP难问题[6]。因此，解决方案大部分集中在启发式算法(heuristic algorithms)中[7]。VNE的过程通常被分为两个阶段：节点映射(node mapping)阶段和链路映射(link mapping)阶段。大部分启发式算法使用基于分别解决两个阶段的节点排序方法[8]-[12]。节点重要性指标通常被使用在节点映射阶段，对虚拟节点和物理节点进行排序。然后在链路映射阶段使用Dijkstra或者多商品流(MCF，Multi-Commodity Flow)算法映射链路。然而，这种缺乏全面考虑的两阶段映射方法没有把节点映射阶段和链路映射的关系考虑进去。同时，这些算法的性能特别依赖节点重要性指标的设计。

强化学习是机器学习的一个领域，使自我改进的代理能够从环境中获得的相互影响中学习，以实现目标[13]。强化学习问题通常表示为马尔科夫决策(MDP)，它的奖励延迟机制适用于延迟敏感问题，比如VNE。因为强化学习属于无监督学习，我们能设置VNE结果的利润，不需要手工标记大量数据。VNE问题的环境是完全可观测的、确定的、离散的，因此对我们来说，很容易对VNE环境建模。我们没有像大多数的启发式方法分别求解两部分映射阶段，而是制定了一个基于MDP的两阶段方法来关联两个阶段。蒙特卡洛树搜索作为一种解决带有大状态空间MDP的方法，不需要像其他搜索树比如A\*或者IDA\*确切的评价函数[14]。能使用在轻赛或者重赛，代替轻赛中的随机移动，在重赛中特定领域的知识可能会被应用在各种启发式算法中来影响移动的选择[15]。

目前，使用MCTS[16]-[18]解决VNE问题的研究者仍然是少部分。我们企图将MCTS和VNE问题的特定领域的知识相结合，因为我们认为好的知识将会指导搜索更好的方向。我们发现，在基于节点排序的方法中使用的节点重要性对结合特定的领域VNE问题知识是一种有效的方法，而且[16]-[18]中没有一种方法考虑过节点重要性。因此，在本文中，我们不再测试哪个节点的重要性度量表现得更好，而是关注如何将节点重要性与基本MCTS相结合的问题。本文提出了两种利用节点重要性的方法：在使用MCTS前确定虚拟节点的映射顺序；MCTS过程中确定将要选择的物理节点的概率。对于虚拟节点，我们使用一个简单的节点重要性度量，该度量应用本地网络资源的知识。具体来说，我们将虚拟节点的重要性定义为节点的CPU请求和所有邻近的链路带宽的总和。这意味着我们将优先考虑映射被认为更重要的虚拟节点，例如，一个请求更多的CPU和链路带宽资源。这有助于避免由于大量映射顺序造成的太多的时间复杂度，并且如果当前情况下没有可行的解决方案，则映射过程将较早地失败，这也有助于加快搜索速度。在启动MCTS之前，我们只计算每个虚拟网络的虚拟节点重要性一次。对于物理节点，设计了一种称为最小带宽消耗的节点重要度度量，这应用物理网络拓扑结构的知识。如果所选择的物理节点在理论上导致带宽的消耗较少，那么节点将被认为更重要，并且在MCTS的扩展和仿真阶段，应该优先选择该节点。

VNE问题的奖励不同于象棋等双人围棋。这是因为VNE的问题不仅仅是输赢。它更适合被视为一个单人游戏，其中一个VNE代理试图最大化映射虚拟网络请求的收入。VNE问题中的回报是成功映射的收益，它可能会是一个比较大的数而不是总是保持在-1或1中。因此，我们建议采用单人MCTS而不是基本MCTS。在选择阶段用SP-UCT[19]( Single-player UCT)替换传统UCT（应用于树的上限置信区间1）来实现这一点。我们认为在SP-UCT中模拟结果的附加方差将引导代理寻找更好的方向。

然后我们将描述我们解决VNE问题的方法。我们首先将VNE问题转化为一个基于模型的强化学习（RL）问题，将映射过程定义为一个有限的MDP，在这个有限的MDP中，我们在模拟环境中运行映射以获得回报。在MDP的每个状态中，我们选择一个物理节点并将一个虚拟节点映射到它上面。在终端状态中，我们使用最短路径算法映射链路，并且返回链路映射成本作为奖励的一部分。由于MDP的奖励延迟机制，链路映射的结果会影响节点映射阶段，使两个嵌入阶段协调。这将改善基于节点排序方法中VNE问题的两个相关映射阶段分离所带来的性能差的问题，在MDP之前，我们必须确定虚拟节点的映射顺序，而不是使用随机顺序。然而，时间复杂度将会很高。因此，我们将节点重要性度量应用于按非递增顺序排列虚拟节点。这就意味着我们将优先考虑映射那些被认为更重要的虚拟节点。之后，我们使用单人游戏的MCTS方法解决基于MDP的协调两阶段的问题。物理节点的节点重要性也应用在我们的单人游戏MCTS的虚拟网络映射在VNE领域特定知识的指导下来扩展树节点，更准确地估计了仿真值。我们的预测结果与MaVEn-S做了比较，这个方法用到了基础MCTS和最短路径链路映射方法来解决VNE问题。本篇论文的主要贡献点是：

(1) 考虑到VNE问题更适合被视为一个单人游戏，试图最大化它的收入，而不是仅仅考虑输赢，我们建议将SP-UCT（19）应用于VNE问题中，以实现开发与探索之间的平衡。

(2) 我们提出了两种应用节点重要性的基本MCTS 方法：一种是在搜索之前对虚拟节点进行排序。在VNE的MCTS方法中，我们应该确定虚拟节点的映射顺序。不是随机地对虚拟节点的顺序进行排序，我们建议根据在经典节点排序方法中使用的节点重要性度量对虚拟节点进行排序。它实质上是“树搜索中的变量排序”。另一种是在这些搜索过程中对物理节点进行排序，在MCTS的模拟和扩展阶段对物理节点的排序类似于“树搜索中的值排序”，代理将优先选择被认为更重要的物理节点。

本文的其余部分安排如下。在第二节中，我们回顾了关于VNE问题的相关工作。在第三节中，我们建立了VNE问题的数学模型，包括网络模型和问题形式化。第四节为VNE问题提出了一种基于MDP方法协调两阶段的方法，第五节提出了一种结合节点重要性知识的单人Monte Carlo树搜索（SP-MCTS）方法来求解VNE问题的MDP。在第六节中，我们验证并分析了我们提出的算法的性能。最后，第七节得出结论。

2.相关工作

VNE的主要目的是在物理网络资源有限的前提下最大化资源利用率[3]。根据数学模型和方法，VNE的优化能分成三类：精确算法，启发式算法，元启发式算法。

精确方法能得到VNE问题的最优解[6]，具体的VNE精确方法[4]、[5]通常由整数线性规划（ILP）和混合整数规划(MIP)等优化理论来表示。例如，Chowdhury等人通过引入元节点，将VNE表示成MIP，并放宽整数规划，得到一个可以在多项式时间内求解的LP表达方法。之后，提出了两种新的VN映射算法，分别是R-ViNE (随机虚拟网络映射) and D-ViNE (确定性虚拟网络映射)。然而，由于巨大的计算复杂性，这种精确方法可能在大规模网络中不常用。

启发式方法在大规模网络中能找到比较好的解，在效率和性能之间做出权衡 [7]。大部分启发式方法[8]-[10]在节点映射阶段应用了节点重要性指标来排序节点，使用最短路径或者MCF算法映射链路。受Google搜索引擎PageRank的启发，向成等人[9]设计了一种基于资源和拓扑属性的拓扑感知节点排序算法NodeRank。两种VNE算法被提出：RW-MaxMatch根据排序将虚拟节点映射到物理节点，使用最短路径或者MCF方法映射链路；RW-BFS是一个基于广度优先搜索的回溯虚拟网络映射算法。

元启发式方法比如粒子群优化[21]、蚁群优化[22]、遗传算法[23]、模拟退火算法[24]，能找到接近最优的解，并且通过避开局部最优来提高性能。比如，[25]的作者，结合离散粒子群优化（DPSO）、模拟退火算法和禁忌搜索技术，解决提前收敛问题。之后就提出了基于混合粒子群优化的VNE算法。

近些年，一些工作已经开始使用强化学习的算法来解决VNE问题，[16]-[18]，[26]，[27]。在[26]中提出了一种基于分散Q-learning的状态值近似算法，用于选择资源分配的动作。目前，使用MCTS的研究还是少数。[16]中Haeri等人使用基本MCTS解决VNE问题的MDP(马尔科夫决策过程)，提出了两个算法，MaVEn-S 和MaVEn-M。MaVEn-S在链路映射阶段用到了最短路径算法，MaVEn-M用到了多商品流算法。[27]中Yao等人提出了一种基于动态属性矩阵表示的强化学习算法，算法中物理节点被表示为静态矩阵，可以动态更新。[18]中，Haipeng Yao等人将mcts与策略网络相结合，进行节点映射决策。他们用VNRs的历史映射方案训练策略网络，然后应用策略梯度自动实现最优映射。我们发现他们当中没有将MCTS和节点重要性知识结合，节点重要性知识广泛应用在节点排序方法中。很多特定领域的知识能被用在预测虚拟节点和物理节点重要性上。因此，通过节点重要性和MCTS结合是一种有效的方式。论文中在扩展和仿真阶段将MCTS和节点重要性结合，来提高搜索效率。

3.虚拟网络映射问题

这部分，介绍VNE问题的数学模型，包括基本网络模型、问题形式化、目标函数和性能度量。

A.网络模型

(1)物理网络(SN)

物理网络被建模成一个加权无向图Gs=(Ns,Es,As),，Ns是物理节点，Es物理链路，As是物理网络的属性。

论文中，我们忽略了节点的内存、节点的存储、链路的延迟等依赖关系，只考虑了CPU容量和链路带宽。物理节点ns有CPU容量Ascpu(ns) ，物理节点i与j之间的物理链路es有带宽Abw，Ps表示物理节点i与j之间的所有的路径。

(2)虚拟网络请求

与物理网络类似，虚拟网络被建模成无向图，有N、E、A，N是虚拟节点Acpu表示虚拟节点的CPU请求，Abw表示虚拟链路的带宽请求。除此之外虚拟网络请求被定义为φ(Gv,td)