

基于增强学习的5G网络切片资源动态优化方案

Dynamic Resources Optimization for 5G Network Slicing Based on Reinforcement Learning

任语铮/REN Yuzheng
谢人超/XIE Renchao
黄韬/HUANG Tao

(北京邮电大学, 北京 100876)
(Beijing University of Posts and
Telecommunications, Beijing 100876, China)

中图分类号: TN929.5 文献标志码: A 文章编号: 1009-6868 (2018) 01-0031-006

摘要: 提出了一种基于增强学习的网络切片资源动态优化方案。使用该方案动态调整网络切片资源时,通过考虑未来网络切片中的业务流量变化情况,对业务流量进行预测,从而推断出未来网络资源的划分情况;再通过增强学习算法,使得未来时刻的网络资源划分状态对当前划分策略做出影响,从而得到当前的最佳策略。基于该算法,可以保证在资源分配过程中对网络需求变化做出快速响应,并通过仿真进行了验证。

关键词: 5G; 网络切片; 增强学习; 动态优化

Abstract: In this paper, a dynamic optimization algorithm based on reinforcement learning for network slicing division is proposed. Network resources can be dynamically allocated in the following ways: the traffic flow can be predicted by considering the changes of flow, then the division of future network resources can be deduced; based on reinforcement learning algorithm, the current partition strategy will be affected by the state of network resource partitioning in the future, and the best division strategy can be got. Based on this algorithm, the change of network requirements can be rapidly responded in the process of resource allocation, and verified by simulation.

Key words: 5G; network slicing; reinforcement learning; dynamic optimization

1 5G 网络架构与网络切片的概念

为了能对各业务提供独立的网络服务而又不铺设专用网络,5G网络引入了网络切片技术^[1-2],即使用软件网络定义(SDN)/网络功能虚拟化(NFV)技术,将物理基础设施资源虚拟化为多个相互独立的平行的网络切片,每个网络切片服务于某一具体的业务场景,以满足不同业务场景对带宽、时延、服务质量等差异化要求,从而满足各种垂直行业多样化需求,以增强网络弹性和自适应性。网络切片技术提升了网络资源利用率,节省了运营商的花费。

下一代移动通信网(NGMN)5G白皮书中的网络切片如图1^[3]所示。目前NGMN对不同的应用场景进行了划分,并为每个应用场景设计了相应的网络切片,总共定义了8个系列,覆盖了现今业务的大部分场景^[3]。

然而网络切片资源的划分并非是一成不变的,而应随各业务流量的

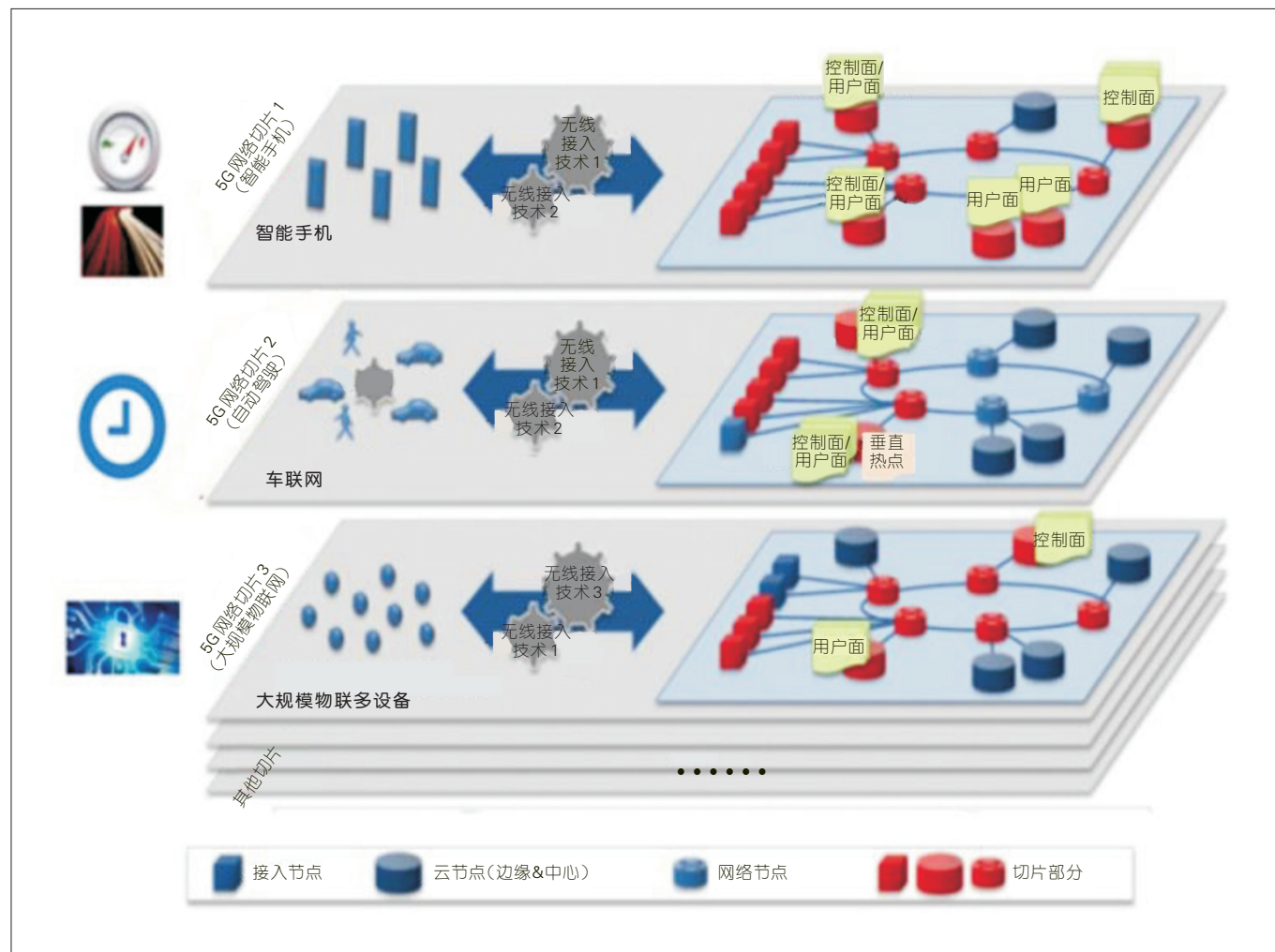
变化进行动态调整。首先,由于部分用户在不断移动,且用户需求往往会随着时间的改变发生改变,所以网络中各业务流量也是动态变化的。此外,第三方服务提供商可能会开发不同的网络业务,形成不同的应用需求,从而要求一个不同的网络切片实例。这都导致了网络切片的划分要随需求的改变发生变动。在基于切片的网络架构中,切片划分的优劣程度直接影响了网络性能,所以如何对切片资源进行动态优化至关重要。

针对网络切片资源的动态优化,有学者提出了一种基于比例公平算法的半静态资源分配方案。该方案

使各网络切片之间能实现更公平地资源分配^[4]。然而该算法更着眼于公平性而不是性能,所以其资源利用率还有提升空间。还有学者认为可以通过对流量进行统计分析,从而得到全网的流量分布特征,再根据流量分布预先构造好基本切片。之后通过分析实时流量的负载和需求构造切片,并将构造结果通过OpenFlow协议下发到交换节点上^[5]。

然而,以上的切片划分算法都是依据当前时刻流量进行优化,而未考虑未来网络流量变化的影响。事实上,在动态优化网络切片资源时需要将未来网络流量的情况一并考虑进

收稿日期: 2017-12-12
网络出版日期: 2018-01-19



▲ 图1 网络切片逻辑架构

来,因为如果在决策时考虑了未来网络的流量变化,则相当于在切片划分策略中引入了预测功能。使得划分结果可以更快地对未来网络的需求变化作出响应。

为了解决这一问题,我们提出了一种基于增强学习的动态优化网络切片资源方案。

2 基于增强学习的网络切片动态优化方案

2.1 增强学习

增强学习(RL)是近些年机器学习和智能控制领域的主要方法之一。RL通过使累积效用最大化,从而确定智能体在环境中应该采取的

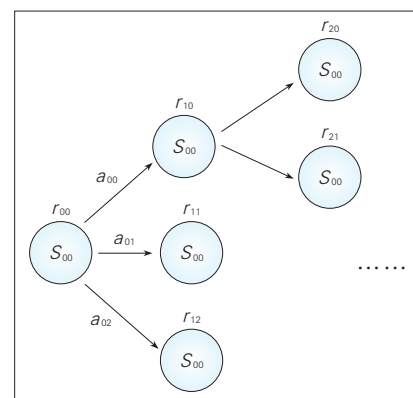
一组行为。RL累积效用的计算方式并不取决于过去的行为,而取决于未来的状态,也就是说未来所处的状态会影响到当前的状态选择。通过增强学习,一个智能体可以知道在某个特定状态下应该采取什么行动。

RL的思想很像马尔可夫过程(MDP)。它定义了四元组 $\{S, A, P_{sa}, R\}$ 。其中, S 是智能体当前所处的状态, A 是智能体采取的行为, P_{sa} 是智能体在状态 S 下做出动作 A 之后转移到其它状态的概率分布, R 是每个状态的效用函数。此外,RL还定义了状态到行为的映射, $\pi: S \rightarrow A$,被称为策略。

如图2所示,状态 S_{00} 在动作 a_{00} 下,以一定的转移概率转移到了 S_{10} ,

状态 S_{00} 的效用函数是 r_{00} ,状态 S_{10} 的效用函数是 r_{10} 。

增强学习通过定义和最优化值函数来得到最优策略,最常见的值函数形式如式(1):



▲ 图2 马尔可夫过程

$$V^\pi(s) = E_\pi \left(\sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s \right) \quad (1)$$

可以看到:这是对一组效用函数的加权和求期望,其中 γ 被称为折合因子,描述了未来效用对当前效用的重要程度。

有了值函数的定义之后,求最优策略就变成了最大化值函数,即:

$$\pi^* = \arg \max_{\pi} V^\pi(s), \forall s \in S \quad (2)$$

我们可以依据以下原则来对已有策略进行改进:若保持策略 π 的其他行为不变,仅将状态 s 下的动作 a 改变为 a' ,由此得到了新策略 π' 。若值函数 $V' > V$,则说明策略 π 好于策略 π' 。我们可以通过动态规划的算法来得到最优策略 π^* 。

2.2 基于增强学习的网络切片动态优化方案

2.2.1 算法的基本思想

端到端的网络切片主要包括无线接入网(RAN)侧切片、核心网(CN)切片,以及连接两者的传输网络切片。本算法主要针对CN切片。在5G网络中,核心网节点分布式协作,以拥有更优的网络性能。

算法的基本思想是:通过定义链路状态矩阵和节点状态矩阵来描述CN切片的不同划分方式,则切片在不同划分方式间的改变就映射成了一张状态转移图,再通过定义各状态下的效用函数和搜索最大化效用函数的方式,得到下一个最佳转移状态,也就得到了CN切片的最优化分方式。

2.2.2 算法中几个重要定义

(1)定义链路状态和节点状态。首先需要对CN资源进行离散化,即对核心网链路与核心网节点资源划分资源片。在动态优化时,以一个资源片为最小的变动单位。链路和节点在划分切片时的不同划分方式对应着不同的状态向量。

(2)预测未来链路和节点所处的状态。如果动态优化算法具有一定的预测功能,那么CN的划分结果则能更快地对网络需求变化作出响应,所以在决策时需要考虑将来时刻核心网可能处于的划分状态。由此我们还需要进行流量分析,从而预测出未来各时刻CN所处的状态。

(3)链路效用函数、节点效用函数与总效用函数。效用函数描述了关心的若干指标。由于对链路和节点往往关心不同的指标,需要分别定义链路效用函数和节点效用函数。

以链路效用函数为例,假设在一个网络中有 n 类业务,则可以定义某状态下的链路效用函数为:

$$V_l = \sum_{i=0}^n V_{li}, \text{ 其中 } V_{li} \text{ 为该链路的第 } i \text{ 个}$$

网络切片的效用函数。第 i 个网络切片的子效用函数可以按照如下的思路定义:首先,链路的利用率应该为一个合适的值,因为链路利用率过高会带来拥塞、丢包;链路利用率过低会带来资源的浪费,所以在定义效用函数时,可以给定一个参考链路利用率 α ,当链路的实际利用率偏离 α 的程度越小时,该网络切片的子效用函数越高。另外,对某一业务而言,不同切片的重要性可能不同,所以可以定义越重要的切片对应的效用函数越高等。

同理,可以定义节点的效用函数 V_n 。于是,在某时刻 t 下,可以定义总的效用函数 $V_t = V_l + V_n$,则 t_0 时刻的最大化目标函数为:

$$V_{\text{总}} = v_0 + \gamma^1 \cdot V_{t_1} + \gamma^2 \cdot V_{t_2} + \dots + \gamma^T \cdot V_{t_T} = \sum_{k=0}^T \gamma^k \cdot V_{t_k} \quad (3)$$

其中, γ 为折合因子,描述了未来状态对当前决策的重要程度。

2.3 算法的具体步骤

算法包括输入和输出。输入指各链路、节点在一段时间内的历史数据包,数据包信息主要包括:数据的业务类型及其对应的网络切片号、数据包长度、时间戳等。

输出指此时应该对网络链路和网络节点做出怎样的划分。

算法的流程如下:

(1)内容预测。基于过去一段时间内核心网中的数据请求情况,预测未来各离散时刻下每条链路中各类业务包的流量情况。

(2)资源离散化。将各链路和网络节点的资源划分资源片,日后在进行资源配置时都以一个资源片为最小的划分单位。

(3)维护两个状态矩阵,分别描述核心网的链路状态和节点状态。假设网络中共有 n 类切片,给定链路初始状态 S_l 。此时第 i 条链路的资源划分情况为 $l_i = (a_1, a_2, \dots, a_n)$,其中, a_k 是第 k 个网络切片分得的资源片数, $\sum_{i=0}^n a_i$ 为该条链路的资源片总数,则可得到整个网络的链路资源划分矩阵 $A_{m \times n}$,其中每个行向量描述了各条链路的资源划分情况。定义 $S_l = A_{m \times n}$ (m 条链路)。

给定节点初始状态 S_n ,此时对第 j 个节点的资源划分情况为 $n_j = (b_1, b_2, \dots, b_n)$,其中 b_k 是第 k 个网络切片分得的资源片数, $\sum_{i=0}^n b_i$ 为该网络节点的资源片总数,则可得到整个网络的节点资源划分矩阵 $B_{k \times n}$,其中每个行向量描述了各节点的资源划分情况。定义 $S_n = B_{k \times n}$ (k 个节点)。

(4)定义效用函数。根据关心的指标定义链路效用函数 V_l 和网络节点效用函数 V_n ;从而得到总的效用函数 $V_t = V_l + V_n$ 。假设关心未来 T 个时刻网络状态对当前决策的影响,则需要最大化的目标函数为:

$$V_{\text{总}} = V_0 + \gamma^1 \cdot V_{t_1} + \gamma^2 \cdot V_{t_2} + \dots + \gamma^T \cdot V_{t_T} = \sum_{k=0}^T \gamma^k \cdot V_{t_k} \quad (4)$$

(5)假设需要优化的为 t_0 时刻,此时的链路状态为 S_l 、节点状态为 S_n ,则可求得相应的链路效用函数和节点效用函数。在某一action下,链路状态和节点状态在 t_1 时刻将转

移到其相邻状态 S'_l 和 S'_n 。 S'_l 和 S'_n 均为集合,里面的元素是当前状态的相邻状态。

由于核心网的链路资源往往比节点资源更为丰富,且不同的业务对网络节点的资源需求不同,例如:有的CN切片需要较大的计算资源,有的切片则需要更大的存储资源,所以每次转移到新的一组状态 S'_l 和 S'_n 之后需要将 S'_l 代入 S'_n 内进行检验,考查新的链路资源划分是否满足此时的节点划分需求,若满足,则可求得新状态下 S'_l 和 S'_n 的效用函数,然后继续进行下一次状态转移;若不满足,则将该状态对标记为无效状态,即从状态转移图中将该转移节点删除。链路状态 S'_l 回退到状态 S_l ,重新转移到 S'_l 集合中的其他状态,再继续进行验证,于是可以得到未来 T 个时刻的网络状态,再根据预测的数据包流量情况算出其各自的效用函数,并以一定的折合率影响当前决策,得到了该组策略对应的效应值:

$$V_{\text{总}} = V_{t_0} + \gamma^1 \cdot V_{t_1} + \gamma^2 \cdot V_{t_2} + \dots + \gamma^T \cdot V_{t_T} = \sum_{k=0}^T \gamma^k \cdot V_{t_k} \quad (5)$$

(6) 最大化目标函数 $V_{\text{总}}$ 。使用动态规划算法进行策略改进,从而收敛到最优策略,即可得到当前时刻与未来时刻的链路和节点的资源划分。

3 仿真结果分析

我们对比考察了3种算法的性能:方法1是基于比例的网络切片动态优化方案,该方案会依据当前各业务流量按比例划分切片资源;方法2是基于公平的静态网络切片分配方案,该方案将网络资源公平地分配到各切片上,且根据网络流量变化进行调整;方法3是基于增强学习的5G网络切片资源动态优化方案,即文中第3部分所述方案。下面的实验模拟5G核心网,其中的核心网节点假设采用分布式协作。首先保证核心网的网络资源不变,通过提升网络中请求数据量,比较3种算法的资源利

用率;再保证核心网资源和请求数据量不变,通过改变折合因子 γ 的取值,考察网络资源的利用率。

假设有4个核心网网关节点,每个节点可提供的资源数按照以40为中心的均匀分布生成,CN节点的资源包括计算资源、存储资源等,两个CN节点间的链路以一定的概率 P 生成,每条链路的资源数按照以55为中心的均匀分布生成。假设运营商同时构建了2个CN切片,各切片的请求数据包数量在一定的范围内随机生成,例如:当各切片的最大内容请求数为 N ,则每条链路各切片请求数均为 $0 \sim N$ 内的随机数,依次可以得到每条链路的数据包请求数,进而可得到通过每个节点的数据包请求数,其值为连接在该节点上的各链路数据量之和的 $1/2$,总共可以生成 T 组数据,用以模拟未来 T 时间内各切片的需求情况。

3.1 网络中数据量的影响

图3、图4描述了5G核心网中,当前时刻各算法的节点平均资源利用率和链路平均资源利用率。节点平均资源利用率依据以下方式得到:首先考察一个CN节点,并根据当前时刻的流量情况算出该节点各切片的节点资源利用率,再根据各切片的资源利用率算出该节点的资源利用率。同理可得到每个CN节点的资源利用率,进而可得到节点的平均资源利用率,以及链路平均资源利用率。图3、图4主要用于考察3种算法对网络资源的利用情况。

图5、图6描述了5G核心网中,未来时刻各算法的节点平均资源利用率和链路平均资源利用率。节点平均资源利用率按照以下方式得到:保持网络划分结果不变,首先考察一个CN节点,根据未来时刻的流量情况算出该节点各切片的资源利用率;再根据各切片的资源利用率算出该节点的平均资源利用率;同理可得到每个CN节点的资源利用率,进而可

得到节点的平均资源利用率,以及链路平均资源利用率。图5、图6主要用于考察各算法的网络划分结果对未来时刻数据流量的匹配程度。

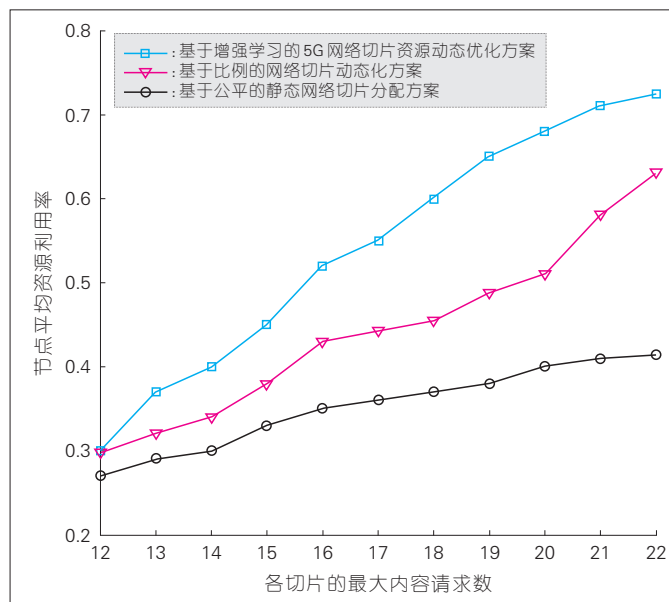
由图3、图4可以看到:当各CN切片的最大内容请求数上升,即核心网中的数据总量上升时,3种算法的资源利用率随之上升,并且两种动态调整算法的资源利用率提升幅度要高于静态划分算法。这主要是因为动态的资源分配结果与网络需求更为匹配,所以让核心网资源被更好地利用。

由图5、图6可以看到:基于增强学习的5G网络切片资源动态优化方案的平均资源利用率最高,基于公平的静态划分算法其次,基于比例的划分算法最后,其中基于比例的网络切片动态优化方案最不稳定。直观地,各CN切片的最大内容请求数增加主要会带来两方面的变化:第一,使得核心网中的数据总量上升,从而资源被更充分的利用;第二,各CN切片流量需求的变动幅度加大。这意味着当网络中流量变动时,基于比例的划分算法需要反复调整网络划分结果,而核心网资源的不断调整会消耗大量时间和资源,所以其资源划分结果对未来网络的需求变化响应较慢,且不具有稳定性。因此当网络划分结果与未来时刻流量不匹配时,会导致资源利用率很低。

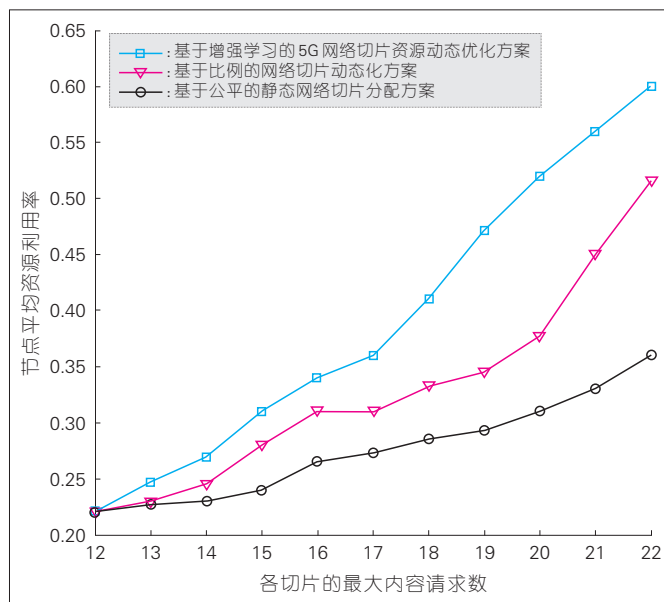
3.2 γ 取值的影响

折合因子 γ 描述了未来效用对当前效用的重要程度。下面我们考察其对基于增强学习的5G网络切片动态优化方案性能的影响。

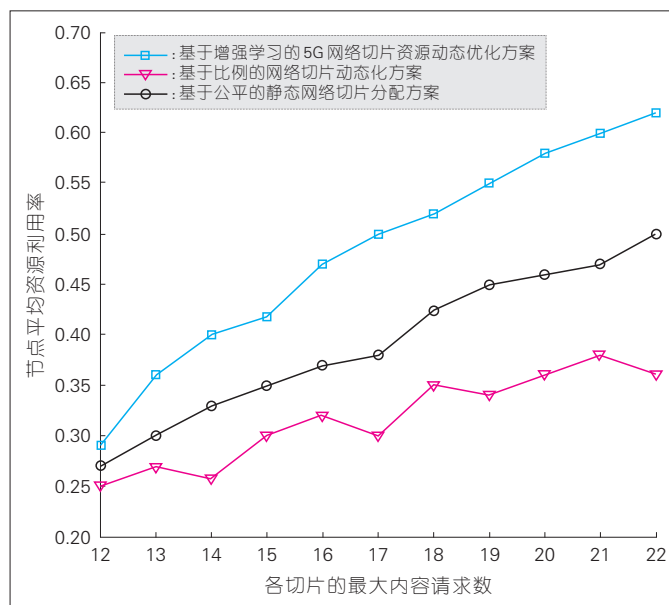
假设运营商同时构建了2个CN切片,拓扑及核心网资源的生成方式与3.1中一致,各切片最大内容请求数为18。共生成 T 组数据,用以模拟未来 T 时间内各切片的需求情况。固定以上参数不变,通过改变折合因子 γ 的取值,考察其对资源利用率的影响,结果如图7、图8所示。节点资



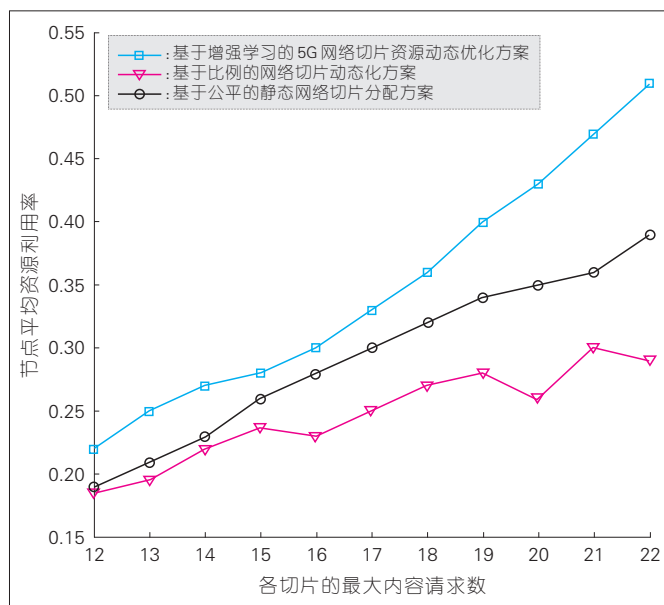
▲ 图3 当前时刻各算法的节点资源利用率



▲ 图4 当前时刻各算法的链路资源利用率



▲ 图5 未来时刻各算法的节点资源利用率



▲ 图6 未来时刻各算法的链路资源利用率

源利用率根据以下方式得到：先考察一个CN节点。分别根据当前时刻、未来时刻的流量情况算出该节点各切片资源利用率；再根据各切片的资源利用率算出该节点的平均资源利用率；同理可得到每个CN节点的资源利用率，进而可得到节点的平均资源利用率，同理可得到链路平均资源利用率。 t_0 为当前时刻的资源利用率， t_1 为未来时刻的资源利用率。

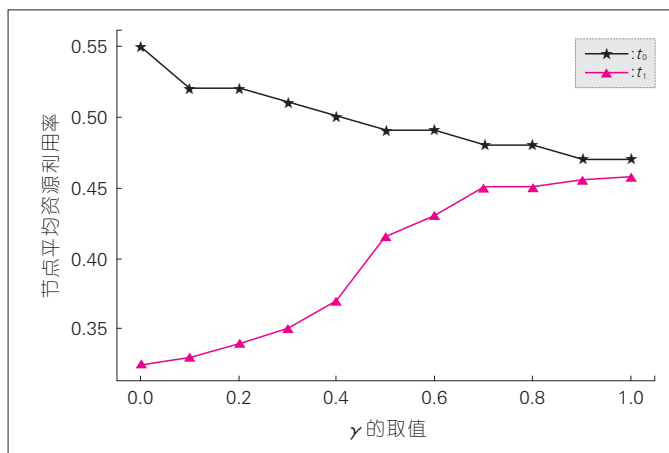
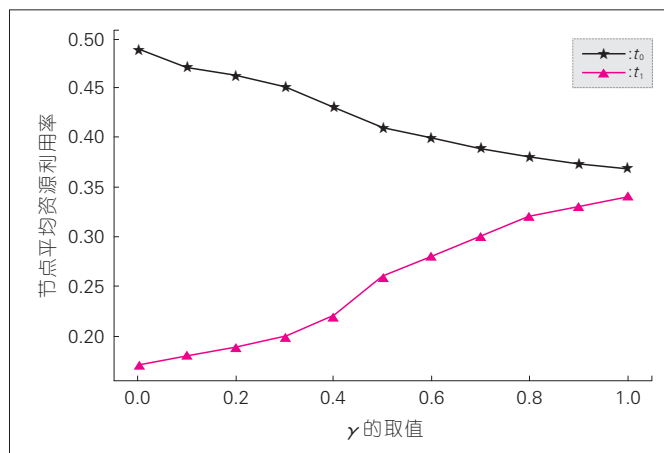
随着折合因子 γ 的增加，当前时刻的资源利用率略有下降，未来时刻的资源利用率大幅提升。

4 算法的简化与参数选择

从前文我们可以看到：增强学习得到的是5G核心网在未来一段时间内的一组划分策略。这组策略除了包括当前策略及未来策略，这意味着我们可以通过检测核心网中数据包的

实际情况来简化计算量。如果未来时刻核心网数据包的实际流量情况与预测的结果相差不大，那么只需要在计算好的未来策略的基础上做演近式微调，就可以得到新的策略。基于这种方式，我们可以大大简化该算法的计算量。

另外，我们还可以通过5G核心网中的链路与节点资源的约束来定义无效状态，并以此来简化计算量。

▲图7 γ 取值对节点资源利用率的影响▲图8 γ 取值对链路资源利用率的影响

因为实际的5G网络中往往会存在链路和节点资源不匹配的问题,所以链路和网络节点则存在着一系列的约束关系。在划网络资源时,移动运营商往往需要根据网络切片类型和网络负载情况为各个网络切片分配固定的资源以满足其最低要求,这说明有些状态是不能被转移的。所以,我们可以定义无效状态,并且在实际计算的过程中不断标注无效状态,删除不符合资源约束条件的状态组合,不断缩小状态转移图的规模,以此来简化计算量。

我们需要考虑的另一个问题是如何对核心网资源进行离散化。网络资源的离散化粒度决定了计算量的大小,如果离散化粒度过小,则状态矩阵维度太高,计算时间太长;如果离散化粒度过大,虽然可以简化计算,但计算结果却很难逼近最优解。另外,每一条链路和网络节点的离散化粒度可以不同。

此外,我们还需要找到合适的折扣因子 γ ,它表明了未来效用对当前效用的影响,并据此来影响当前的决策。 γ 的值不可过大或者过小,如果 γ 过大,而对数据包的情况预测与事实相差很大,则会导致对将来时刻的效用计算不够准确,进而影响当前策略的选择;如果 γ 过小,则将来时刻的状态对此刻选择的策略影响不够,就可能达不到我们的目的,无法更好

地反应各切片对资源的需求变化。

最后我们还需要选择合适的的时间窗 T ,它描述了我们考虑将来多长时间内的状态与效用函数。可以看到: T 不可以过大,一方面 T 值的选择直接决定了本文算法的计算量大小;另一方面,如果 T 值过大,而对数据包的情况预测与事实相差很大,则会导致对将来时刻的效用计算不够准确,进而影响当前策略的选择。但 T 也不可以过小,如果 T 太小则对未来状态考虑得不够,同样不能更好地实现我们的目的。

5 结束语

针对如何动态调整5G核心网网络切片资源,我们提出了一种基于增强学习的5G网络切片资源动态优化方案。该算法首先对未来时刻的核心网状态进行预测,再依据核心网未来时刻所处的状态对当前决策做出影响,以使切片划分结果更快地对网络的需求变化做出响应、改善网络整体性能。

该算法中的一些参数会较大地影响性能,如何对这些参数进行选取在日后还值得更多的研究。

参考文献

- [1] ITRI. Updates to Solution 6.1.3: Introducing Network Instance ID: S2-162666[R/OL]. (2016-08-30)[2016-08-30].http://

www.3gpp.org/ftp/tsg_sa/wg2_arch/tsgs2_115_Nanjing_China/Docs/

- [2] IMT-2020 (5G)推进组.5G网络架构设计白皮书[R/OL].(2016-08-30)[2016-08-30].http://www.catt.cn/kxyj/qwfb/bps/201606/t2016024_2173031.html

- [3] NGMN Alliance. 5G White Paper[R]. 2015

- [4] 栗欣,龚金金,曾捷.面向5G网络切片无线资源分配[J].电子产品世界,2017,24(04):30-32+40

- [5] 周恒,畅志贤,杨武军,郭娟.一种5G网络切片的编排算法[J].电信科学,2017,33(08):130-137

作者简介



任语铮,北京邮电大学在读硕士研究生;主要研究方向为基于人工智能的未来网络体系架构,现从事基于人工智能的视频缓存与调度算法研究。



谢人超,北京邮电大学副教授、硕士生导师;主要研究方向为未来网络体系架构、信息中心网络与边缘计算等;主持国家与省部级项目4项;已发表论文40余篇,申请授权专利10余项。



黄韬,北京邮电大学教授、博士生导师;主要研究方向为未来网络体系架构、软件定义网络、信息中心网络等;主持国家与省部级项目10余项;已发表论文100余篇,申请授权专利40余项。