

运营商网络中基于深度强化学习的服务功能链迁移机制

陈卓^{*①②} 冯钢^② 何颖^② 周杨^③

^①(重庆理工大学计算机科学与工程学院 重庆 200433)

^②(电子科技大学通信抗干扰技术国家级重点实验室 成都 710077)

^③(奥本大学计算机科学与软件工程学院 美国阿拉巴马州 奥本市 36849)

摘要: 为改善运营商网络提供的移动服务体验, 该文研究服务功能链(SFC)的在线迁移问题。首先基于马尔可夫决策过程(MDP)对服务功能链中的多个虚拟网络功能(VNF)在运营商网络中的驻留位置迁移进行模型化分析。通过将强化学习和深度神经网络相结合提出一种基于双深度Q网络(double DQN)的服务功能链迁移机制, 该迁移方法能在连续时间下进行服务功能链的在线迁移决策并避免求解过程中的过度估计。实验结果表明, 该文所提出的策略相比于固定部署算法和贪心算法在端到端时延和网络系统收益等方面优势明显, 有助于运营商改善服务体验和资源的使用效率。

关键词: 运营商网络; 迁移机制; 深度强化学习; 服务功能链

中图分类号: TN915; TP393

文献标识码: A

文章编号: 1009-5896(2020)09-2173-07

DOI: 10.11999/JEIT190545

Deep Reinforcement Learning Based Migration Mechanism for Service Function Chain in Operator Networks

CHEN Zhuo^{*①②} FENG Gang^② HE Ying^② ZHOU Yang^③

^①(College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 200433, China)

^②(National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 710077, China)

^③(Department of Computer Science and Software Engineering, Auburn University, Auburn 36849, United States of America)

Abstract: To improve the service experience provided by the operator network, this paper studies the online migration of Service Function Chain(SFC). Based on the Markov Decision Process(MDP), modeling analysis is performed on the migration of multiple Virtual Network Functions(VNF) in SFC. By combining reinforcement learning and deep neural networks, a double Deep Q-Network(double DQN) based service function chain migration mechanism is proposed. This method can make online migration decisions and avoid over-estimation. Experimental result shows that when compared with the fixed deployment algorithm and the greedy algorithm, the double DQN based SFC migration mechanism has obvious advantages in end-to-end delay and network system revenue, which can help the mobile operator to improve the quality of experience and the efficiency of resources usage.

Key words: Operator network; Migration mechanism; Deep Reinforcement Learning (DRL); Service Function Chain (SFC)

1 引言

运用网络功能虚拟化(Network Function

Virtualization, NFV)^[1]技术可以使软件化的虚拟网络功能(Virtualized Network Function, VNF)更高效地部署在通用硬件平台而非特定专用硬件设备, 为运营商优化网络资源部署带来了更优的灵活性。在5G/B5G核心网中, 根据不同业务将所需的VNF逻辑连接成服务功能链(Service Function Chain, SFC)^[2,3]是提供适应性服务最有效的方式。但由于各类移动应用在位置动态性和业务流时变性的特点

收稿日期: 2019-07-18; 改回日期: 2020-06-14; 网络出版: 2020-07-14

*通信作者: 陈卓 chen_zhuo@cqu.edu.cn

基金项目: 国家自然科学基金(61471089, 61401076)

Foundation Items: The National Natural Science Foundation of China (61471089, 61401076)

突出,使得运营商基于SFC提供持续可靠的服务面临极大挑战。运营商按照固定方式完成SFC部署后,一方面会使业务的服务质量难以得到持续有效保证,另外导致IT资源使用效率低下。因此针对业务的时变特性,适时地迁移SFC之上VNF的部署位置以实现SFC的动态调配显得尤其重要。学界已经关注到该问题带来的各种性能影响^[4-8],但仍缺乏有针对性的解决方案。其中,文献[4]综合考虑SFC提供服务的过程中链路和通用服务节点使用之间的关系,给出了VNF部署和路径选择策略。文献[5]在预定义的用户移动模型前提下给出了内容在多个云数据中心中进行迁移决策机制,该机制综合考虑了迁移代价和用户体验。Eramo等人^[6]提出通过VNF部署迁移和SFC路径选择,减少被拒绝的SFC请求。文献[7]以提高网络的整体收益为目标,研究并提出了在面对网络负载动态变化的情况下的VNF资源扩展策略。相比之下,本文把VNF所驻留的虚拟机(Virtual Machine, VM)位置对迁移带来的影响因素纳入考虑,有效地保证迁移后的端到端数据传输延迟。从研究方法来看,深度强化学习通过最大化“智能体”(agent)从“环境”中获得的累计奖赏值,通过自学习的方式达成目标的最优策略^[8]。因此适合于在“环境”变化的应用场景中,让“智能体”学习解决问题的策略。在类似研究中,文献[9]和文献[10]在软件定义网络环境下以深度强化学习方法为基础,分别研究了适应性路由问题和多媒体业务流的流量控制。与已有工作相比,本文则在运营商网络中考虑了SFC中多个VNF之间的逻辑关系并研究SFC的迁移问题。

本文研究基于NFV构建的运营商网络中SFC的迁移问题,针对业务的动态性和移动性对SFC迁移重配置问题进行形式化描述,并进一步建立了马尔可夫决策过程加以分析,基于深度强化学习方法双深度Q网络(double Deep Q-Network, double DQN),提出了一种面向运营商网络的SFC在线迁移机制。通过实验的对比分析表明本文所提算法能够通过优化网络的资源配置,提高移动业务的服务质量,使网络系统收益得到明显改善。本文第2节描述形式化分析模型,第3节提出基于double DQN的SFC在线迁移机制,第4节建立实验场景并和已知的算法策略进行性能对比。

2 服务功能链迁移的形式化描述

2.1 优化目标

在基于NFV的运营商网络中,假设为某移动业务提供服务的SFC上有 N 个按序逻辑链接起来的VNF,即: v_1, v_2, \dots, v_N ,这 N 个VNF组成集合 V 。其中 v_1 为该SFC中距离入口网元最近的VNF,

v_N 为该SFC中距离出口网元最近的VNF。移动业务的移动性和时变性需要SFC做在线迁移的时候,网络决策机构(如:网络控制器)可以考虑的迁移方案集合中的元素为:迁移SFC中的 C_N^1 个VNF,迁移SFC中的 C_N^2 个VNF, ..., 迁移SFC中的 C_N^N 个VNF,也即迁移方案集合有 $\sum_{i=1}^N C_N^i$ 种可行的迁移方案。但考虑SFC中多个VNF之间的逻辑关联关系,可将迁移方案集合精简为包括了 $N+1$ 种迁移方案的集合,该集合包括:不迁移VNF、单独迁移VNF v_1 ,同时迁移VNF v_1 和 v_2 ,同时迁移VNF v_1, v_2 和 v_3, \dots ,迁移SFC之上的所有VNF v_1, v_2, \dots, v_N 。定义 $X_{1,2}(t)$ 表示VNF v_1 和 v_2 之间的传输距离, $X_{i,i+1}(t)$ 表示VNF v_i 和 v_{i+1} 之间的距离,所以每段链路的数据传输时延可表示为式(1),其中 C 为速率

$$d_i(t) = \frac{X_{i-1,i}(t)}{C} \quad (1)$$

由于每迁移1个VNF将在运营商网络中新启用一个通用服务节点,假设新启用一个通用服务节点并实例化为VNF的时延为 d_m ,则迁移 i 个VNF的迁移时延为 $i \cdot d_m$ 。对于网络决策机构的一种迁移方案,需要迁移 K 个VNF,得到迁移时延为 $i \cdot d_m + d_N$,可表示为

$$d_K = K \cdot d_m + \frac{\sum_{n=1}^{N-1} X_{n,n+1}(t)}{C} \quad (2)$$

本文目标是通过SFC之上VNF的合理迁移,使得业务在移动过程中平均时延最小,引入了折扣因子 η^t 表达时间对期望的影响, η^t 随着时间的增加而减小,时间越久影响则越小,最终得到优化目标为

$$\min E \left[\sum_{t=0}^{\infty} (\eta^t \cdot d_K) \right] \quad (3)$$

2.2 系统模型

若将上述整个运营商网络作为一个系统加以分析,当前时刻 t 的系统状态只与 $t-1$ 时刻的系统状态有关,而与 $t-1$ 时刻以前系统状态无关。考虑到系统具有马尔可夫性以及业务的时变特性,本文基于连续时间马尔可夫决策过程(Markov Decision Process, MDP)对系统进行形式化分析。MDP是指智能体(agent)根据连续观察系统的状态,从可用的行动集中选用一个行动作为行动决策,使系统转移到下一个状态并产生回报,然后再根据新观察到的系统状态再做出新的决策,反复进行以获得系统的长期最优受益。一个马尔可夫决策过程可以描述

为 $M = \{S, (A(i), i \in S), P, R, \eta\}$ 其中：(1) S 为所有系统状态集合，用于描述当前情况的所有参数；(2) $(A(i), i \in S)$ 为可能出现的动作(action)的集合；(3) P 为状态转移概率矩阵，即不同状态转移之间的概率关系；(4) R 为回报函数，描述来自环境的反馈，回报可以是正或负。回报为正对应奖励，为负对应惩罚；(5) η 为折扣因子， $\eta \in [0, 1]$ ，另外定义智能体的策略为： $\pi(a|s) = P[A_t = a|S_t = s]$ 。

根据马尔可夫决策过程的描述，并结合本文所研究的运营商网络中SFC的在线迁移问题，可做如下定义。(1)状态集合：系统所有可能的状态集合。把每段SFC链路的传输时延作为状态，每次业务的时变性导致的迁移都会使部分链路发生变化，即产生状态之间的转换。将此状态集合表示为 $S^t = \{y_1(t), y_2(t), \dots, y_n(t)\}$ 其中， $y_i(t)$ 表示第 i 个VNF部署在节点 $y_i(t)$ 上。另外， $x_i(t)$ 表示VNF v_i 到VNF v_{i+1} 之间的传输距离，该距离最终用于计算累加回报。(2)动作空间：本文把运营商网络中SFC迁移的 $N+1$ 种潜在可能的动作定义为动作空间，即一种SFC迁移方案是动作集中的一个元素。其中，不做SFC的迁移定义为动作 a_0^t ，只迁移SFC中和入口网元网络距离最短的一个VNF v_1 对应动作 a_1^t ，SFC中迁移距离入口网元网络距离由小至大的前 i 个VNF v_1, v_2, \dots, v_i 对应动作 a_i^t ，而迁移SFC中所有VNF v_1, v_2, \dots, v_N 对应动作 a_N^t 。因此动作集合为： $a^t = \{a_0^t, a_1^t, \dots, a_i^t, \dots, a_N^t\}$ ， a_i^t 表示迁移SFC中距离入口网元网络距离由小至大的前 i 个VNF ($0 \leq i \leq N$)，而SFC中VNF序列中的后 $N-i$ 个VNF不做迁移。(3)效益模型：本文定义了系统的奖励和惩罚，把奖励作为系统收入，把惩罚作为系统支出。由传统模型能得到端到端总时延，而为了表示业务实际时延以及业务能容忍的最大端到端时延 D 之间的关系，若用户端到端时延 $d_t \leq D$ ，表明该任务能成功完成。否则任务失败，接受惩罚。(4)定义惩罚函数为： $F(d_t) = d_t - D$ 迁移VNF会带来代价，例如：IT资源占用量的增加。因此把迁移SFC中的VNF数量的因素引入到系统的惩罚中，如式(4)所示

$$y(m) = \sum_{i=1}^N f_i \quad (4)$$

其中 $f_i=1$ 表示迁移VNF v_i ，而 $f_i=0$ 表示不迁移VNF v_i 。所以基于系统的状态和动作，SFC迁移模型的效益函数可表示为

$$R_t(s^t, a^t, s^{t+1}) = i(s^t, a^t, s^{t+1}) - \beta \cdot e(s^t, a^t, s^{t+1}) \quad (5)$$

式(5)中 $i(s^t, a^t, s^{t+1})$ 是系统在状态为 s^t 选择行动 a^t 后，系统所获得的总收益。在本研究问题中，我们

用移动业务的满意度来表示系统总收益，而满意度基于Sigmoidal函数^[1]来衡量，表示为

$$i(s^t, a^t, s^{t+1}) = 1 - \exp \left(-\frac{\omega_2 \cdot \frac{1}{d_t^2}}{\omega_1 + \frac{1}{d_t}} \right) \quad (6)$$

其中， $d_t = i \cdot d_m + \frac{1}{c} \sum_{i=1}^{N-1} x_i(t)$ 表示端到端时延， ω_1 和 ω_2 是用于调节Sigmoidal幅度的参数。另外，式(5)中 $\beta \cdot e(s^t, a^t, s^{t+1})$ 表示系统的总支出， β 为惩罚因子， $e(s^t, a^t, s^{t+1})$ 是系统提供的服务未达到业务满意度时的总开支，表示为 $e(s^t, a^t, s^{t+1}) = F(d_t) \cdot f(s^t, a^t)$ 。当 $d_t - D \geq 0$ 时， $f(s^t, a^t) = y(m)$ ，当 $d_t - D < 0$ ， $f(s^t, a^t) = 1/y(m)$ 。由于 $e(s^t, a^t, s^{t+1})$ 的引入，运营商网络中的网络控制器在进行SFC迁移决策的时候，会同时将移动业务满意度和迁移代价纳入考虑。因此所选择的行动以提升系统整体回报为导向，避免了系统只考虑满意度的改善而频繁的对SFC中的VNF进行迁移重部署。本文的目标是得到一个优化策略 π^* ，即在相应的状态 s^t 下采用动作 a^t 后，使效益函数最大化，即求解式(7)的优化问题

$$\pi^* = \arg \max_{\pi} E \left[\sum_{t=0}^{\infty} \eta^t \cdot R_t(s^t, s^{t+1}, a^t) \right] \quad (7)$$

其中， η^t 为折扣因子 ($0 \leq \eta \leq 1$)，并且 η^t 随着时间增加其值减少。得到最优策略 π^* 为一系列系统中SFC中的VNF迁移动作。

3 基于神经网络的深度Q学习方法

3.1 强化学习基本原理

强化学习(Reinforcement Learning, RL)模型主要描述智能体以试错的机制与环境进行反复交互，通过最大化累积回报的方式来学习最优策略。基于RL策略的模型如图1所示，该模型由5个关键部分组成，包括：状态 S 、动作 a 、状态转移概率 P 、回报 r 以及策略 $\pi(s, a)$ 。在智能体(agent)和环境交互的过程中，智能体在不同的时间点根据观察状态与系统回报，根据策略 $\pi(s, a)$ 执行相应的动作(action)。在行动之后，智能体的状态以概率 P 的描述转换到下一状态，同时智能体获得来自环境的反馈回报 r 。由于智能体的当前状态影响下一状

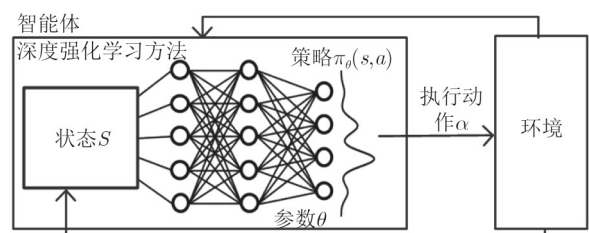


图1 基于神经网络的强化学习决策

态, 而和当前状态之前的状态无关, 因此可以用第2.3节的MDP来描述强化学习模型。

通过RL建模, 其核心是能够得到 $\pi(s, a): S \times A \rightarrow [0, 1]$, 也即: 得到智能体的状态空间和动作空间到概率的映射。通常智能体的状态空间和动作空间巨大, 这要求RL方法能够利用有限学习经验, 记忆完成大范围空间有效知识的获取与表示。当运营商网络规模足够大的情况下, 系统状态空间的规模将使得解方程组难以实现。更关键的是, SFC迁移模型的状态转移概率矩阵无法提前获得, 这使得经典的策略迭代法和值迭代法都难以使用。在近期的研究工作中, 深度神经网络(Deep Neural Network, DNN)被成功的用于求解强化学习模型^[9,12]并得到了较好的结果。

3.2 基于双DQN的SFC在线迁移机制

在充当“智能体”的网络决策机构处于某一种状态时可以选择多种动作, 而不同动作的执行则让智能体进入下一个不同的状态。本文引入了动作价值函数 $Q^\pi(s, a)$ 来估计每个动作的价值, 将式(5) $R_t(s^t, a^t, s^{t+1})$ 简记为 r_t , 由此, 动作价值函数表示为 $Q^\pi(s, a) = E[r_{t+1} + \eta r_{t+2} + \eta^2 r_{t+3} + \dots | s, a]$, 进一步表示为

$$Q^\pi(s, a) = E_{s'}[r + \eta Q^\pi(s', a') | s, a] \quad (8)$$

为了得到最优策略, 需要求解最优动作价值函数即

$$Q^*(s, a) = E_{s'}[r + \eta \max_{a'} Q^*(s', a') | s, a] \quad (9)$$

值迭代算法就是根据更新 Q 值来使其收敛到最优, Q -学习的思想就是完全根据值迭代得到的。但值迭代每次都要把所有的 Q 值更新, 但针对本文所研究的SFC迁移问题, 难以遍历整个状态空间, 因此 Q -学习只使用有限的样本进行 Q 值更新

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \eta \max_{a'} Q^*(s', a') - Q(s, a)] \quad (10)$$

虽然根据值迭代算法可以得到目标 Q 值, 但并没有将得到的 Q 值赋予新的 Q 值, 而是采用了渐进的方式不断逼近目标, 类似梯度下降。而式(10)中的 α 表示控制之前 Q 值和新 Q 值差别的学习率, 可以减少误差。 η 表示折损率, 即: 未来的经验对当前状态执行动作的重要程度。最后可以收敛到一个最优的 Q 值。本文采用深度 Q 学习算法解决连续状态问题。深度 Q 网络(DQN)^[13]是结合 Q 学习的深度神经网络, DQN将深度学习模型和强化学习相结合并能直接从高维的输入学习控制策略, DQN能解决多重维度的复杂动作问题。本文引入值函数

$Q(s, a; \psi)$ 表示输入任意状态得到输出, 其目的是将 Q 值的复杂更新问题转换成一个函数问题, 相似状态对应相似的动作实现值函数的近似, 然后继续用 $Q(s, a; \psi) \approx Q^*(s, a)$ 加以表示, 其中参数 ψ 表示神经网络的权重。DQN通过更新参数 ψ 来使近似 Q 函数无限逼近最优的 Q 值, 使其转化成函数优化问题。由于函数的非线性特性, 本文采用深度神经网络来作为这个近似 Q 函数, 也即采用深度强化学习的方法, 并基于此提出一种基于双DQN的SFC迁移的方法, 有效避免了过度估计带来的影响。双DQN将目标 Q 值中选择和评估动作分离, 让他们使用不同的 Q 函数(网络)。其中一个用于产生贪婪策略, 另一个用来产生 Q 函数估计值, 因此实现需要两个 Q 函数网络, 原DQN的 Q 函数网络被称为在线网络, 后者称为目标网络。双DQN算法使用的目标可用式(11)表达

$$Y_t^{\text{doubleQ}} = r_{t+1} + \gamma Q(s_{t+1}, \arg \max_{a'} Q(s_{t+1}, a; \psi_t^-); \psi_t^-) \quad (11)$$

双DQN中计算目标使用了两个不同的 ψ_t , 分别来自当前 Q 网络以及目标 Q 网络, 当前 Q 网络负责选择动作, 带有时延 ψ_t^- 的目标 Q 网络负责计算目标 Q 值。另外, 使用经验池来解决相关性及非静态分布的问题, 经验池将每个时间步智能体与环境交互得到的转移样本 (s^t, a^t, r^t, s^{t+1}) 存储到回放记忆单元, 需要训练时, 随机拿出一部分调参样本进行训练, 这样的随机抽取克服了相关性问题。基于双DQN的SFC迁移方法流程可通过图2加以描述。

基于双DQN的算法优势在于能通过 Q 学习构造损失函数, 然后通过经验池(experience replay)解决相关性和非静态分布问题, 同时能使用目标网络解决稳定性问题。表1描述了SFC迁移机制的伪代码。

4 实验与性能评估

4.1 实验场景建立

本节对所提出的基于双DQN的SFC迁移机制进行数值仿真。为评估其性能, 将其与常采用的SFC固定部署机制^[4,14]以及基于贪心策略(greedy)的SFC迁移机制^[15]作性能对比。其中, SFC固定部署机制在业务的移动和网络时变情况下, SFC中的各VNF部署位置均不发生改变, 该机制将不会带来网络额外开销。而基于贪心策略的SFC迁移机制中, 在决策VNF部署位置的变更时, 总选择系统收益最大的加以执行。本文选择端到端的业务流时延和网络系统收益作为关键的评估指标, 其中端到端的业务流时延是影响移动业务体验的最关键因

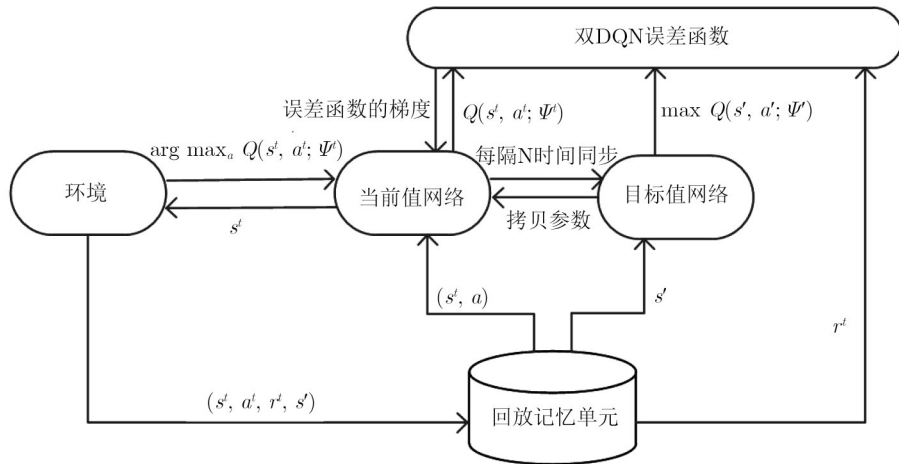


图 2 基于双DQN的SFC迁移方法流程图

表 1 基于双DQN的SFC迁移算法的伪码

输入：运营商网络拓扑 $G = (N, E)$ ，服务功能链集合 C ，网络功能集合 F ；
输出：SFC迁移策略；
步骤1：初始化随机权重为 ψ 的神经网络；
步骤2：初始化动作值函数 Q ；
步骤3：初始化经验池(experience replay)存储器 N ；
步骤4：for episode = 1, 2, ..., M do,
观察初始状态 s^0 ，
for $t = 0, 1, \dots, N-1$ do,
以概率为 ϵ 选择一个随机动作 a^t ，
否则选择动作 $a^t = \arg \max Q(s^t, a; \psi^t)$ ；
在仿真器中执行动作 a^t ，并观察回报 R_{t+1} 和新状态 s_{t+1} ，
存储中间量 $\langle s^t, a^t, r^t, s_{t+1} \rangle$ 到经验池存储器 N 中，
从经验池存储器 N 中获取一组样本，
计算损失函数 $L(\psi^t)$ ，
计算关于 ψ^t 的损失函数的梯度，
更新 $\psi^t \leftarrow \psi^t - \phi \nabla_{\psi^t} L(\psi^t)$ ，其中 ϕ 为学习率；
end
end

素，因此重点对其进行评估。另外，网络系统收益反映了不同策略对于网络资源的占用情况，也作为重要指标加以评估。仿真平台基于Matlab搭建，拓扑的建立参考文献[16]中经典的NSF-NET，NSF-NET网络包含14个通用物理节点用于承载VNF所需的计算资源，同时包括23条物理网络链路，多条SFC可共享物理网络链路。

4.2 实验结果与分析

首先定义包括5个VNF的SFC，行动空间中的迁移方案数为6。随着移动业务移动次数的增加，端到端时延发生变化，采用基于双DQN的SFC迁

移机制与不迁移SFC的端到端延迟对比如图3所示。该图表明当移动业务位置发生改变且移动前后的业务流经由不同的入口网元进入运营商网络时，基于双DQN的SFC迁移机制对端到端时延保障有明显效果。在已发生200次的业务流时变情况下，相比于固定的SFC部署方式平均降低了38.5%的端到端延迟，有效改善了业务体验。另外，定义了3种SFC，分别包括3个、5个和7个VNF的情况下，行动空间中迁移方案数分别为4、6和8的情况下，基于双DQN方法的端到端的延迟。由图4可以看到，业务的时变特性导致SFC上的VNF发生迁移，根据业务移动前后的入口网元的位置差异，为改善端到端延迟而选择了不同的迁移方案。可以看到，当业务需要更多的VNF提供服务时，由于迁移方案涉及到更多数量的VNF，因此导致平均延迟更高。图4中，当SFC包括7个VNF的时候，分别比5个VNF和3个VNF的移动业务延迟平均提高了97.4%和41.7%。

系统收益基于式(8)描述，其调节参数设定为 $\omega_1=0.9$ 和 $\omega_2=0.4$ 时，当用户随机移动200次的过程中，本文比较了双DQN迁移机制、greedy迁移机制以及不迁移SFC情况下的系统收益，结果如图5

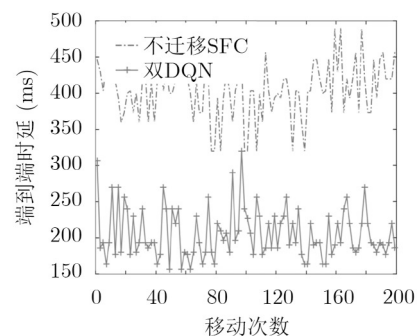


图 3 移动业务端到端时延对比

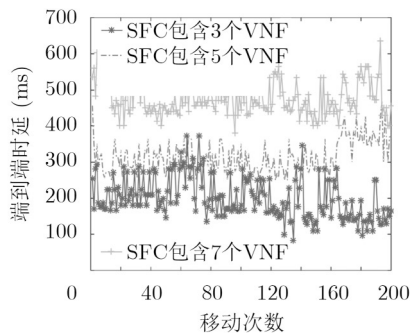


图4 基于双DQN方法在不同长度VNF下的延迟对比

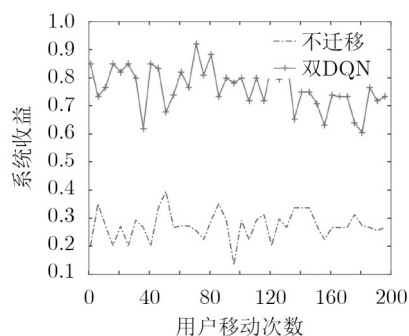


图5 不同移动次数下系统收益对比

所示。可以看到,通过对SFC之上的VNF进行合理的位置调整后,整个运营商网络的系统收益有明显提高,平均比不迁移SFC机制的系统收益提升为142%。本文提出的迁移机制在系统收益上得到了显著改善。同时,我们对基于双DQN的SFC迁移机制、基于贪心策略的迁移机制(greedy迁移)以及不迁移SFC情况下的系统累积收益进行了对比,结果如图6所示,该图统计了10000次迭代过程中的系统累积收益,由于固定方式部署SFC不迁移的情况下,没有适时调整VNF的部署位置会引起某时间段系统收益下降,导致系统累积收益增加不明显。而基于双DQN的SFC迁移机制能实现在多个时间点的连续决策,同时有效避免决策的局部最优,相比于基于greedy迁移机制能更明显的提升系统累积收益。该图说明了基于双DQN的SFC迁移机制在

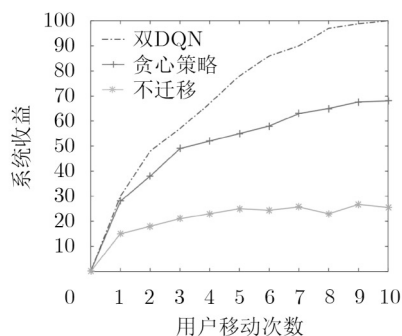


图6 系统累积收益对比

改善业务流的服务延迟的同时能有效改善系统资源的占用。

5 结束语

本文深入研究了运营商网络中服务功能链的迁移问题。首先,分析了时变特性明显的移动业务难以在固定的SFC部署机制下接受持续满意的服务,并提出对SFC中的VNF进行迁移的必要性。然后,本文基于马尔可夫决策过程(MDP)对迁移问题进行建模,定义了相关的关键参数,状态空间、动作空间与系统回报。进一步,本文把深度Q学习与神经网络相结合提出了一种基于双DQN的SFC迁移机制。最后,本文通过仿真实验,将提出的迁移机制和其它相关算法进行对比。结果表明,本文提出的SFC迁移方案在端到端时延和网络系统收益等方面得到了明显改善。本文提出的策略有助于运营商规划设计更优化的网络决策机构,改善各种移动新业务的服务体验。

参考文献

- [1] CHATRAS B and OZOG F F. Network functions virtualization: The portability challenge[J]. *IEEE Network*, 2016, 30(4): 4-8. doi: 10.1109/MNET.2016.7513857.
- [2] ZHANG Qixia, LIU Fangming, and ZENG Chaobing. Adaptive interference-aware VNF placement for service-customized 5G network slices[C]. *IEEE Conference on Computer Communications*, Paris, France, 2019: 2449-2457. doi: 10.1109/INFOCOM.2019.8737660.
- [3] AGARWAL S, MALANDRINO F, CHIASSERINI C F, et al. Joint VNF placement and CPU allocation in 5G[C]. *IEEE Conference on Computer Communications*, Honolulu, USA, 2018: 1943-1951. doi: 10.1109/INFOCOM.2018.8485943.
- [4] KUO T W, LIOU B H, LIN K C J, et al. Deploying chains of virtual network functions: On the relation between link and server usage[C]. *The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, USA, 2016: 1-9. doi: 10.1109/INFOCOM.2016.7524565.
- [5] TALEB T, KSENTINI A, and FRANGOUDIS P A. Follow-me cloud: When cloud services follow mobile users[J]. *IEEE Transactions on Cloud Computing*, 2019, 7(2): 369-382. doi: 10.1109/TCC.2016.2525987.
- [6] ERAMO V, MIUCCI E, AMMAR M, et al. An approach for service function chain routing and virtual function network instance migration in network function virtualization architectures[J]. *IEEE/ACM Transactions on Networking*, 2017, 25(4): 2008-2025. doi: 10.1109/TNET.2017.2668470.
- [7] HOUIDI O, SOUALAH O, LOUATI W, et al. An efficient algorithm for virtual network function scaling[C]. 2017

- IEEE Global Communications Conference, Singapore, 2017: 1–7. doi: 10.1109/GLOCOM.2017.8254727.
- [8] CHO D, TAHERI J, ZOMAYA A Y, *et al.* Real-time Virtual Network Function (VNF) migration toward low network latency in cloud environments[C]. The 10th IEEE International Conference on Cloud Computing, Honolulu, USA, 2017: 798–801. doi: 10.1109/CLOUD.2017.118.
- [9] 兰巨龙, 于倡和, 胡宇翔, 等. 基于深度增强学习的软件定义网络路由优化机制[J]. 电子与信息学报, 2019, 41(11): 2669–2674. doi: 10.11999/JEIT180870.
- LAN Julong, YU Changhe, HU Yuxiang, *et al.* A SDN routing optimization mechanism based on deep reinforcement learning[J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2669–2674. doi: 10.11999/JEIT180870.
- [10] HUANG Xiaohong, YUAN Tingting, QIAO Guanhua, *et al.* Deep reinforcement learning for multimedia traffic control in software defined networking[J]. *IEEE Network*, 2018, 32(6): 35–41. doi: 10.1109/MNET.2018.1800097.
- [11] LEE J W, MAZUMDAR R R, and SHROFF N B. Non-Convex optimization and rate control for multi-class services in the Internet[J]. *IEEE/ACM Transactions on Networking*, 2005, 13(4): 827–840. doi: 10.1109/TNET.2005.852876.
- [12] 李晨溪, 曹雷, 陈希亮, 等. 基于云推理模型的深度强化学习探索策略研究[J]. 电子与信息学报, 2018, 40(1): 244–248. doi: 10.11999/JEIT170347.
- LI Chenxi, CAO Lei, CHEN Xiliang, *et al.* Cloud reasoning model-based exploration for deep reinforcement learning[J]. *Journal of Electronics & Information Technology*, 2018, 40(1): 244–248. doi: 10.11999/JEIT170347.
- [13] MNIH V, KAVUKCUOGLU K, SILVER D, *et al.* Human-level control through deep reinforcement learning[J]. *Nature*, 2015, 518(7540): 529–533. doi: 10.1038/nature14236.
- [14] GHAZNAVI M, KHAN A, SHAHRIAR N, *et al.* Elastic virtual network function placement[C]. The 4th IEEE International Conference on Cloud Networking, Niagara Falls, Canada, 2015: 255–260. doi: 10.1109/CloudNet.2015.7335318.
- [15] SUGISONO K, FUKUOKA A, and YAMAZAKI H. Migration for VNF instances forming service chain[C]. The 7th IEEE International Conference on Cloud Networking, Tokyo, Japan, 2018: 1–3. doi: 10.1109/CloudNet.2018.8549194.
- [16] LIN Tachun, ZHOU Zhili, TORNATORE M, *et al.* Demand-aware network function placement[J]. *Journal of Lightwave Technology*, 2016, 34(11): 2590–2600. doi: 10.1109/JLT.2016.2535401.
- 陈 卓: 男, 1980年生, 副教授, 博士, 硕士生导师, 研究方向为网络虚拟化和软件定义网络.
- 冯 钢: 男, 1964年生, 教授, 博士生导师, 研究方向为计算机网络和无线接入网技术.
- 何 颖: 女, 1993年生, 硕士生, 研究方向网络虚拟化技术.
- 责任编辑: 余 蓉