

# 基于随机学习的接入网服务功能链部署算法

陈前斌\* 杨友超 周钰 赵国繁 唐伦

(重庆邮电大学通信与信息工程学院 重庆 400065)

**摘要:** 针对5G云化接入网场景下物理网络拓扑变化引起的高时延问题, 读文提出一种基于部分观察马尔可夫决策过程(POMDP)部分感知拓扑的接入网服务功能链(SFC)部署方案。该方案考虑在5G接入网C-RAN架构下, 通过心跳包观测机制感知底层物理网络拓扑变化, 由于存在观测误差无法获得全部真实的拓扑情况, 因此采用基于POMDP的部分感知和随机学习而自适应动态调整接入网切片的SFC的部署, 优化SFC在接入网侧的时延。为了解决维度灾问题, 采用基于点的混合启发式值迭代算法求解。仿真结果表明, 该模型可以优化部署接入网侧的SFC, 并提高接入网吞吐量和资源利用率。

**关键词:** 网络切片; SFC动态部署; 网络拓扑感知; 部分观察马尔可夫决策过程

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2019)02-0417-07

DOI: 10.11999/JEIT180310

## Deployment Algorithm of Service Function Chain of Access Network Based on Stochastic Learning

CHEN Qianbin YANG Youchao ZHOU Yu ZHAO Guofan TANG Lun

(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** To solve problem of the high delay caused by the change of physical network topology under the 5G access network C-RAN architecture, this paper proposes a scheme about dynamic deployment of Service Function Chain (SFC) in access network based on Partial Observation Markov Decision Process (POMDP). In this scheme, the system observes changes of the underlying physical network topology through the heartbeat packet observation mechanism. Due to the observation errors, it is impossible to obtain all the real topological conditions. Therefore, by the partial awareness and stochastic learning of POMDP, the system dynamically adjust the deployment of the SFC in the slice of the access network when topology changes, so as to optimize the delay. Finally, point-based hybrid heuristic value iteration algorithm is used to find SFC deployment strategy. The simulation results show that this model can support to optimize the deployment of SFC in the access network side and improve the access network's throughput and resource utilization.

**Key words:** Network slice; Service Function Chain (SFC) dynamic deployment; Network topology perception; Partial Observation Markov Decision Process (POMDP)

### 1 引言

网络切片(Network Slicing, NS)是指在一个物理网络基础设施上建立若干个相互隔离的逻辑网络<sup>[1-6]</sup>。每个切片网络包含若干条服务功能链(Service Function Chain, SFC), 且SFC由若干有序虚拟网络功能(Virtual Network Function, VNF)组成<sup>[7]</sup>,

如何部署SFC是一个巨大的挑战。

文献[8,9]的SFC部署方案是在核心网(Evolved Packet Core, EPC)一侧, 优化SFC在EPC侧的时延等服务质量。文献[8]中是基于近似马尔可夫模型(Markov Approximately, MA), 该方案是一个静态部署策略, 这种方案无法实时保证服务质量。文献[9]中基于遗传算法(Genetic Algorithm, GA)的带宽资源动态分配模型实现SFC在核心网的动态部署而适应变化的数据到达量, 但该算法局限于物理网络拓扑不变的情况。同时, 由于5G接入网采用新型C-RAN (Cloud-Radio Access Network)架构<sup>[10]</sup>, 上述部署方案无法直接支持接入网切片中SFC部署。

收稿日期: 2018-04-02; 改回日期: 2018-09-03; 网络出版: 2018-09-12

\*通信作者: 陈前斌 chenqb@cqupt.edu.cn

基金项目: 国家自然科学基金(61571073)

Foundation Item: The National Natural Science Foundation of China (61571073)

针对上述存在的问题，本文提出了一种基于POMDP (Partial Observation Markov Decision Process)的接入网切片的SFC动态部署方案。该方案首先考虑5G接入网采用新型CU (Centralized Unit)\DU (Distributed Unit)分离架构，具体设计了接入网侧的SFC部署流程。

其次部分感知接入网底层物理网络的动态变化。通过心跳包检测的方式检测服务器处于正常与否，并分析观测误差。

最后依据部分感知获得网络拓扑变化情况，通过POMDP动态部署SFC，优化系统的总时延。且基于点的混合启发式值迭代算法解决维度灾问题。

## 2 系统模型

### 2.1 系统场景

本文的研究内容是基于5G接入网C-RAN架构场景下网络切片在上行条件下如何动态优化部署接入网侧的SFC。场景如图1所示，CU和DU设备都实现虚拟化，从而形成DU池和CU池，两者之间通过前传网络进行数据通信<sup>[1]</sup>。不同切片可以灵活部署SFC的VNF满足服务质量，如图1中切片1的SFC2

的VNF1部署在DU池，而切片2的SFC1把VNF1和VNF2都部署在DU池，其余在CU池实例化。同时考虑接入网的缓存功能，所以SFC在DU侧存在队列。

在上行条件下的C-RAN架构中，接入网侧的SFC的VNF有多种部署方式，如图2所示。由于接入网SFC中的各VNF之间的数据交互量不同，不同部署方式使SFC对DU和CU之间的前传网络的时延和带宽需求不同。不同的切片方式一方面会影响资源池的资源分配，另一方面影响SFC的前传资源分配。

CU池和DU池基础设施层是通用服务器节点在实际中会受到环境中随机因素的影响而发生故障<sup>[2]</sup>，从而使服务器网络拓扑发生变化。考虑节点失效引起的拓扑变化，在虚拟化层加入了节点监测功能感知拓扑变化，本文的目标是提出基于拓扑感知的接入网切片的SFC动态部署从而优化网络时延。

### 2.2 基于拓扑感知的SFC部署

本文引入心跳包检测机制<sup>[13]</sup>检测节点*i*在时刻*t*的状态 $\varphi_i(t)$ 。节点*i*的阈值 $T_{o,i}$ 的取值会带来不同程度的误判失效的误差，无法获得全部节点的真实状态。这里采用一种基于概率密度的心跳检测算法

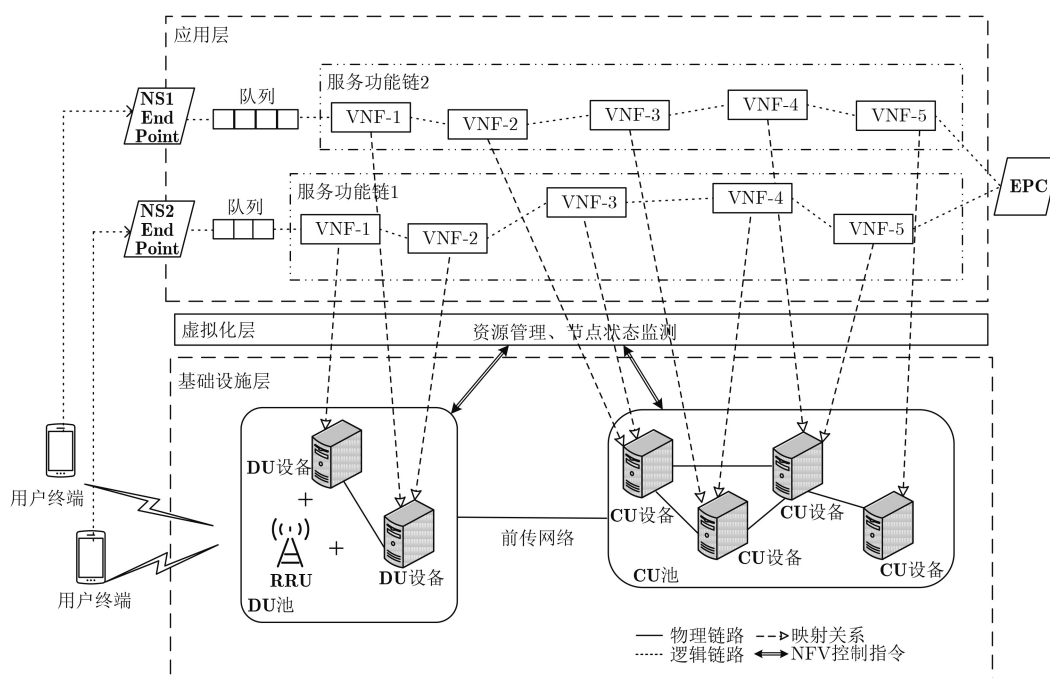


图1 系统模型

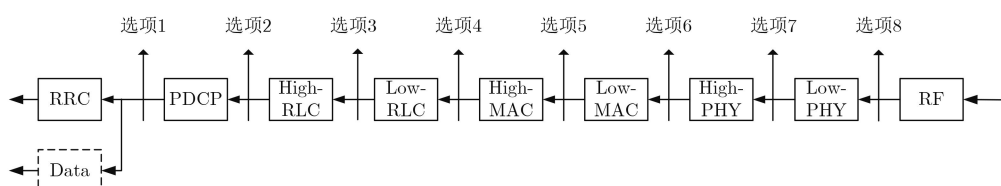


图2 接入网VNF部署方式

分析失效误判概率，本模型假设包到达的间隔时间满足正态分布，设节点*i*阈值为 $T_{o,i}$ ，则节点*i*误判

$$\text{失效的概率 } P_{e,i} = \frac{1}{\sqrt{2\pi}\sigma} \int_{T_{o,i}}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx,$$

$\mu$ 表示样本空间的平均值， $\sigma$ 表示样本空间的标准差。

服务器网络表示为 $G = \{N, L\}$ ， $N = N_1 \cup N_2$ 表示节点集且由DU池的节点集 $N_1$ 和CU池的节点集 $N_2$ 组成， $L = L_1 \cup L_2$ 表示链路集且由DU池的链路集 $L_1$ 和CU池的链路集 $L_2$ 。  $v_{n_k}$ 是服务器节点 $n_k$ 的计算资源容量， $v_l$ 表示链路 $l$ 的带宽容量。设系统的切片集合为 $U$ ， $\ell_u$ 表示切片 $u$ 里的SFC集合。

本文中接入网侧的切片SFC部署流程具体包括3个阶段。第1阶段是系统决策是否修复所有失效的节点， $a_\alpha = 1$ 表示修复所有失效的节点，其余为0。第2阶段是为每个切片网络选择接入网VNF部署方式， $a_\beta(t) = \{\beta_u(t) | u \in U\}$ ， $\beta_u(t)$ 表示切片 $u$ 时刻 $t$ 选择的VNF部署方式，并设 $A_\beta$ 表示所有可能选择的集合。第3阶段分配资源，设当前时刻SFC资源分配方式 $a_\chi(t) = \{\chi_{u,m}^c(t), \chi_{u,m}^b(t) | u \in U, m \in \ell_u\}$ ，将资源离散化，并所有可能的资源分配方式组成集合 $A_\chi$ 。  $\chi_{u,m}^c(t) = \left\{ v_{f_{u,m}^c, n_k}^c \cdot x_{f_{u,m}^c, n_k}^p \mid j \in F_{u,m}, n_k \in G_{f_{u,m}^c}^j \right\}$ 表示 $t$ 时刻SFC  $m$ 的计算资源分配方式，其中 $F_{u,m}$ 表示第 $m$ 个SFC的VNF集合， $G_{f_{u,m}^c}^j$ 表示第 $j$  ( $j \in F_{u,m}$ )个VNF可以实例化的节点集合， $v_{f_{u,m}^c, n_k}^c$ 表示第 $j$ 个VNF在节点 $n_k$ 所得的计算资源， $x_{f_{u,m}^c, n_k}^p = 1$ 表示第 $j$ 个VNF部署在节点 $n_k$ 上，否则为0。  $\chi_{u,m}^b(t)$ 表示 $t$ 时刻SFC  $m$ 的每个VNF的带宽资源分配情况， $\chi_{u,m}^b(t) = \left\{ v_{f_{u,m}^b, l}^b \cdot y_{f_{u,m}^b, l}^p \mid j \in F_{u,m}, l \in L \right\}$ 其中， $v_{f_{u,m}^b, l}^b$ 表示第 $j$ 个VNF在链路 $l$ 上分配的带宽资源， $y_{f_{u,m}^b, l}^p = 1$ 表示第 $j$ 个VNF用链路 $l$ 发送数据，其余情况为0。

### 3 网络模型

#### 3.1 网络优化目标

在时刻 $t$ 基于拓扑感知部署SFC时，一方面会涉及修复失效的节点，会带来修复成本函数 $R_1(t)$ ：

$$R_1(t) = - \frac{\sum_{i \in N_1} c\mu_i(t) + \sum_{j \in N_2} c\mu_j(t)}{(|N_1| + |N_2|)c} \quad (1)$$

其中， $c$ 表示费用单价， $\mu_i = 1$ 表示时刻 $t$ 节点 $i$ 被修复，否则为0， $|N_1|$ 和 $|N_2|$ 分别表示DU池和CU池的服务器数量。

另一方面涉及时延，考虑每条SFC的排队时延和从DU到CU的调度时延，设总时延函数为 $R_2(t)$ ：

$$R_2(t) = - \left[ \sum_{u \in U} \gamma_u \sum_{m \in \ell_u} \left( \frac{q_{u,m}(t)}{\lambda_{u,m}(t)} + \dots \sum_{j \in F_{u,m}}^{n_k} p_{f_{u,m}^j}^{n_k} \right) + \dots \sum_{j \in F_{u,m}} \left( \sum_{l \in L} y_{f_{u,m}^j, l}^p \cdot \frac{D_{f_{u,m}^j}}{v_{f_{u,m}^j, l}^b} \right) \right] / \eta \quad (2)$$

其中， $\gamma_u$ 表示切片 $u$ 的优先级， $\lambda_{u,m}(t)$ 表示SFC  $m$ 的数据到达率，假设SFC的数据到达量满足非齐次泊松分布。 $F_{u,m}'$ 表示SFC  $m$ 的VNF集合且不包括在CU末端的VNF， $D_{f_{u,m}^j}$ 表示第 $j$ 个VNF发送的数据量， $p_{f_{u,m}^j}^{n_k}$ 表示第 $j$ 个VNF在节点 $n_k$ 的处理时延， $\eta$ 是时延归一化参数。接入网切片的SFC部署的优化目标：

$$\max R(t) = e_1 R_1(t) + e_2 R_2(t) \quad (3)$$

其中， $e_1$ 和 $e_2$ 是正值且 $e_1 + e_2 = 1$ 。

#### 3.2 基于POMDP的接入网SFC部署模型

为了在心跳包检测机制感知拓扑变化存在误差的情况下寻找最优的接入网侧的SFC部署方案，本文提出了一个基于POMDP的部署模型，在不确定环境中部分感知拓扑变化情况而决策出最优部署的场景。

##### 状态空间 $S$

$$S = \{(q, \varphi) | \forall q \in Q', \varphi \in \varphi'\}$$

其中， $q$ 为队列状态， $\varphi$ 为拓扑状态。 $Q' = \{(q_{1,1}, \dots, q_{1,|\ell_1|}, \dots, q_{U,|\ell_U|}) | \forall q_{u,m} \in N_Q\}$ 表示队列空间， $q_{u,m}$ 为SFC  $m$ 的队列长度，且队列长度的取值范围 $N_Q = \{0, 1, \dots, n_q\}$ ， $n_q$ 是最大队列长度。

$\varphi' = \{(\varphi_1, \dots, \varphi_{|N_1|}, \dots, \varphi_{|N_1|+|N_2|}) | \varphi_i \in \{0, 1\}\}$ 表示拓扑的状态空间， $\varphi_i = 1$ 表示节点 $i$ 处于正常状态， $\varphi_i = 0$ 表示该节点处于失效状态。

##### 动作空间 $A$

$$A = \{(a_\alpha, a_\beta, a_\chi) | \forall a_\alpha \in \{0, 1\}, a_\beta \in A_\beta, a_\chi \in A_\chi\}$$

其中， $a_\alpha$ 是节点修复动作， $a_\beta$ 是切片选择接入网VNF部署方式的动作， $a_\chi$ 是资源分配动作。

##### 转移概率 $P$

设转移概率为 $\tau(s(t), a(t), s(t+1))$ ， $s(t)$ 表示时刻 $t$ 的系统状态， $a(t)$ 表示 $t$ 时刻采取的动作， $s(t+1)$ 表示 $t+1$ 时刻的系统状态。

$$\tau(s(t), a(t), s(t+1)) = \Pr[q(t+1) | q(t), a(t)] \cdot \Pr[\varphi(t+1) | \varphi(t), a(t)] \quad (4)$$

$q(t)$ 和 $q(t+1)$ 分别表示时刻 $t$ 和 $t+1$ 的队列状态。

$$\begin{aligned} \Pr[q(t+1)|q(t), a(t)] &= \Pr[W_{u,m}(t) \\ &= q_{u,m}(t+1) - q_{u,m}(t) + V_{u,m}(t)] \quad (5) \end{aligned}$$

$W_{u,m}(t)$ 是第 $m$ 个SFC在时刻 $t$ 的数据到达量, 设 $V_{u,m}(t)$ 是时刻 $t$ 的SFC  $m$ 数据服务量, SFC  $m$ 在时刻 $t$ 的队列长度为 $q_{u,m}(t)$ 。 $\varphi(t)$ 和 $\varphi(t+1)$ 分别表示时刻 $t$ 和 $t+1$ 的拓扑状态。

$$\begin{aligned} \Pr[\varphi(t+1)|\varphi(t), a(t)] \\ &= \prod_{i \in N_1} \Pr[\varphi_i(t+1)|\varphi_i(t), a(t)] \\ &\cdot \prod_{j \in N_2} \Pr[\varphi_j(t+1)|\varphi_j(t), a(t)] \quad (6) \end{aligned}$$

如果时刻 $t$ 采取修复所有失效节点的动作, 此时如果 $\varphi_i(t+1) = 1$ , 则 $\Pr[\varphi_i(t+1)|\varphi_i(t), a(t)] = 1$ , 其他为0。不修复失效节点时, 当 $\varphi_i(t) = 0$ , 当且仅当 $\varphi_i(t+1) = 0$ 时 $\Pr[\varphi_i(t+1)|\varphi_i(t), a(t)] = 1$ , 对于当前正常节点受环境影响可能会失效, 即当 $\varphi_i(t) = 1$ 时,

$$\Pr[\varphi_i(t+1)|\varphi_i(t), a(t)] = \begin{cases} p_i, & \varphi_i(t+1) = 0 \\ 1 - p_i, & \varphi_i(t+1) = 1 \end{cases} \quad (7)$$

$p_i$ 表示在随机环境中下一时刻节点 $i$ 失效的概率。

### 回报函数 $R$

设 $t$ 时刻的接入网侧SFC部署动作的回报函数 $R(s(t), a(t))$ 为节点修复成本函数和系统总时延函数的加权和, 即 $R(s(t), a(t)) = e_1 R_1(t) + e_2 R_2(t)$ 。

### 观测空间 $Z$

$Z = \{(z_Q, z_\varphi) | z_Q \in Z'_Q, z_\varphi \in Z'_\varphi\}$ , 其中 $z_Q$ 表示队列观测状态,  $Z'_Q$ 表示队列观察空间, 由于队列状态是完全可知的, 所以 $Z'_Q = Q'$ 。 $z_\varphi$ 表示拓扑观测状态,  $Z'_\varphi$ 表示拓扑观测空间。

### 观测函数 $O$

设 $t+1$ 时刻的观测状态 $z(t+1)$ 的观测函数为 $O(z(t+1), a(t), s(t+1))$ , 其表示为

$$\begin{aligned} O(z(t+1), a(t), s(t+1)) \\ &= O(z_q(t+1), a(t), q(t+1)) \\ &\cdot O(z_\varphi(t+1), a(t), \varphi(t+1)) \quad (8) \end{aligned}$$

其中, 由于队列是完全可知, 当队列观测状态 $z_q(t+1) = q(t+1)$ 时,  $O(z_q(t+1), a(t), q(t+1)) = 1$ , 否则为0。设节点 $i$ 在时刻 $t+1$ 的观测状态为 $z_{\varphi,i}(t+1)$ 时,

$$\begin{aligned} O(z_\varphi(t+1), a(t), \varphi(t+1)) \\ &= \prod_{i \in N_1} \Pr[z_{\varphi,i}(t+1)|a(t), \varphi_i(t)] \\ &\cdot \prod_{j \in N_2} \Pr[z_{\varphi,j}(t+1)|a(t), \varphi_j(t)] \quad (9) \end{aligned}$$

任意节点 $i \in N$ , 如果 $a(t) = (1, a_\beta(t), a_\chi(t))$ , 则节点观测为正常状态, 即 $z_{\varphi,i}(t+1) = 1$ 时,  $\Pr[z_{\varphi,i}(t+1)|a(t), \varphi_i(t+1)] = 1$ , 其他情况为0。如果 $a(t) = (0, a_\beta(t), a_\chi(t))$ , 即当 $\varphi_i(t+1) = 0$ , 仅当 $z_{\varphi,i}(t+1) = 0$ 时 $\Pr[z_{\varphi,i}(t+1)|a(t), \varphi_i(t+1)] = 1$ , 其余情况为0, 而当 $\varphi_i(t+1) = 1$ 时,

$$\begin{aligned} \Pr[z_{\varphi,i}(t+1)|a(t), \varphi_i(t+1)] \\ &= \begin{cases} 1 - P_{e,i}, & z_{\varphi,i}(t+1) = 1 \\ P_{e,i}, & z_{\varphi,i}(t+1) = 0 \end{cases} \quad (10) \end{aligned}$$

### 信念状态

设时刻 $t$ 的信念状态 $\mathbf{b}_t = (b_t(s_1), \dots, b_t(s_{|S|}))$ , 其中 $b_t(s_i)$ 表示系统状态为 $s_i$ 的概率, 且 $\mathbf{b}_t$ 是 $|S|$ 维的向量。基于信念状态执行动作, 获得观察结果, 更新信念状态, 不断循环下去<sup>[14,15]</sup>, 所执行的动作组成一个策略 $\pi$ , 且所有可能的策略的集合为 $\xi$ , 回报函数转化为 $R_b(\mathbf{b}_t, a) = \sum_{s \in S} R(s, a(t)) \cdot b_t(s)$ , 其中 $R_b(\mathbf{b}_t, a)$ 表示信念点 $\mathbf{b}_t$ 的回报函数。基于信念状态, 将接入网切片的SFC部署优化目标函数转化为

$$\max_{\pi \in \xi} \tilde{\rho}\pi = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \gamma^t R_b(\mathbf{b}_t, a(t)) \quad (11)$$

其中,  $\gamma$ 表示折扣因子且 $0 < \gamma \leq 1$ ,  $T$ 表示时间段。

由于该模型中的马尔可夫链是各态历经的, 当 $T \rightarrow \infty$ 时一定存在一个稳定的最优分配策略使目标函数最大<sup>[16]</sup>, 信念点 $\mathbf{b}_t$ 在 $k$ 次迭代后最优策略:

$$\begin{aligned} \pi_k^*(\mathbf{b}_t) &= \arg \max_{a \in A} \left\{ R_b(\mathbf{b}_t, a) \right. \\ &\left. + \gamma \sum_{z \in Z} \Pr[z|\mathbf{b}_t, a] \cdot V_{k-1}^*(\mathbf{b}_{t+1}) \right\} \quad (12) \end{aligned}$$

其中,  $\Pr[z|\mathbf{b}_t, a]$ 是归一化参数,  $V_{k-1}^*(\mathbf{b}_{t+1})$ 表示信念点 $\mathbf{b}_{t+1}$ 在 $k-1$ 次迭代后的最优值函数。

## 4 算法分析和描述

基于点的混合启发式值迭代(Mingled Heuristic Value Iteration, MHVI)算法由于选择值函数上下界较大的点来寻找最远后继信念点更新当前的探索信念点集合 $B_{su}$ , 从而结合信念点的分布情况和值函数选择有效的后继点, 提高了求解效率。该算法在达到阈值条件之前反复执行两个过程: 一是寻找新的信念点来更新 $B_{su}$ , 二是在基于更新后 $B_{su}$ 上执行更新值函数向量集合 $\Gamma$ 。

### 4.1 更新探索信念点

在这个子过程中首先在当前的 $B_{su}$ 中选择值函数上下界差值大于阈值的点构成被扩点集合 $B^{pr}$ :

$$B^{\text{pr}} = \left\{ \mathbf{b} \mid \bar{V}(\mathbf{b}) - \underline{V}(\mathbf{b}) \geq \frac{\varepsilon}{\gamma^{h_b}} \right\} \quad (13)$$

其中,  $h_b$ 表示信念点 $\mathbf{b}$ 的层数,  $\varepsilon$ 是常数。  $\bar{V}(\mathbf{b})$ 表示 $\mathbf{b}$ 处的上界取值, 该值将近似求得,  $\underline{V}(\mathbf{b})$ 表示 $\mathbf{b}$ 处的下界取值,  $\underline{V}(\mathbf{b}) = \max_{\alpha \in \underline{V}} \mathbf{b} \cdot \alpha$ 。接着寻找集合 $B^{\text{pr}}$ 所包含的信念点的后继信念点中上下界差值较大的点组成集合 $\text{su}(\mathbf{b})$ :

$$\text{su}(\mathbf{b}) = \left\{ \mathbf{b}' \mid \mathbf{b}' = \mathbf{b}_a^z, \bar{V}(\mathbf{b}') - \underline{V}(\mathbf{b}') \geq \frac{\varepsilon}{\gamma^{h_{\mathbf{b}'}}} \right\}, \quad \forall a, \forall z \quad (14)$$

从集合 $\text{su}(\mathbf{b})$ 中选取距离 $B_{\text{su}}$ 最远的信念点 $\mathbf{b}''$ :

$$\mathbf{b}'' = \arg \max_{\mathbf{b}'' \in \text{su}(\mathbf{b}), \mathbf{b}'' \notin B_{\text{su}}} \|\mathbf{b}'' - B_{\text{su}}\|_1 \quad (15)$$

其中,  $\|\mathbf{b}'' - B_{\text{su}}\|_1 = \min_{\mathbf{b}^* \in B_{\text{su}}} |\mathbf{b}'' - \mathbf{b}^*|$ , 并将满足条件的后继信念点加入到 $B_{\text{su}}$ 中。

上界集合 $\bar{V}$ 由若干个 $\langle \mathbf{b}_i, \bar{v}_i \rangle$ 组成, 并随着 $B_{\text{su}}$ 更新而更新, 每次更新时新加入点-值对,

$$\begin{aligned} \bar{V}' &= \bar{V} \cup (\mathbf{b}, \max_{a \in A} R_b(\mathbf{b}, a)) \\ &+ \gamma \sum_{z \in Z} P(z | \mathbf{b}, a) \bar{V}(\mathbf{b}_a^z), \mathbf{b} \in B_{\text{su}} \end{aligned} \quad (16)$$

其中,  $\mathbf{b}_a^z$ 表示信念点 $\mathbf{b}$ 在动作 $a$ 和观测 $z$ 的后继点。

下界集合 $\underline{V}$ 由若干个下界向量构成, 更新时先找出点 $\mathbf{b}$ 的新的下界向量 $\alpha_b$ , 并加入集合 $\underline{V}$ 中,

$$\begin{aligned} \alpha_b(s) &= \max_{a \in A} (R(a, \mathbf{b}) + \gamma \sum_{z \in Z} \max_{\alpha \in \underline{V}} \sum_{s' \in S} \tau(s, a, s') \\ &\quad \cdot O(z, a, s') b_a^z(s') \alpha), \\ \alpha_b &= (\alpha_b(s_1), \alpha_b(s_2), \dots, \alpha_b(s_{|S|})) \end{aligned} \quad (17)$$

$b_a^z(s')$ 表示信念点 $\mathbf{b}$ 的后继点中状态为 $s'$ 的概率。

更新探索信念点集合算法示于表1。

#### 4.2 更新值函数向量集

时刻 $t$ 的值函数向量集 $\Gamma_t = \{\alpha_1, \alpha_2, \dots, \alpha_{|\Gamma_t|}\}$ 且最优值函数 $V^*(\mathbf{b}) = \max_{\alpha \in \Gamma_t} \mathbf{b} \cdot \alpha$ 。首先计算动作 $a$ 的一步回报向量 $\Gamma_1^a = (R(s_1, a), R(s_2, a), \dots, R(s_{|S|}, a))$ , 接着需要获得动作 $a$ 在观察 $z$ 的向量集合 $\Gamma_{t+1}^{a,z}$ ,

$$\begin{aligned} \Gamma_{t+1}^{a,z} &= \{\alpha_i^{a,z} \mid \alpha_i^{a,z}(s) = \gamma \sum_{s' \in S} \tau(s, a, s') \\ &\quad \cdot O(z, a, s') \alpha_j'(s'), \alpha_j' \in \Gamma_t, s \in S\} \end{aligned} \quad (18)$$

然后选择使信念点 $\mathbf{b}$ 与 $\Gamma_{t+1}^{a,z}$ 中乘积最大的向量与 $\Gamma_1^a$ 相加获得回报向量 $\Gamma_{t+1, \mathbf{b}}^a$ , 并将每一个动作的回报向量构成集合 $\Gamma_{t+1, \mathbf{b}} = \bigcup_{a \in A} \Gamma_{t+1, \mathbf{b}}^a$ , 在 $\Gamma_{t+1, \mathbf{b}}$ 寻找和 $\mathbf{b}$ 乘积最大的向量 $\text{up}(\mathbf{b}) = \arg \max_{\alpha \in \Gamma_{t+1, \mathbf{b}}} \mathbf{b} \cdot \alpha$ 。

表1 算法1: 更新探索信念点集合 $B_{\text{su}}$

- 
- (1) 用式(13)计算被扩点集 $B^{\text{pr}}$
  - (2) **for all**  $\mathbf{b} \in B^{\text{pr}}$  **do**
  - (3)     用式(14)计算 $\text{su}(\mathbf{b})$
  - (4)     用式(15)计算离 $B_{\text{su}}$ 最远的后继信念点 $\mathbf{b}''$
  - (5) **end for**
  - (6) 清空集合 $V'$ 的元素
  - (7) **for all**  $\mathbf{b} \in B_{\text{su}}$  **do**
  - (8)     用式(17)计算下界向量 $\alpha_b$ 并加入 $V'$ 中
  - (9) **end for**
  - (10) 将下界集合 $\underline{V}$ 更新为 $V'$
  - (11) **for all**  $\mathbf{b} \in B_{\text{su}}$  **do**
  - (12)      $V_{\infty} \leftarrow \{\mathbf{b} \mid \exists s \in S, b(s) = 1\}$
  - (13)      $v_b^0 \leftarrow \sum_{\mathbf{b}' \in V_{\infty}} v(\mathbf{b}') \cdot \mathbf{b}$
  - (14)     **for all**  $\langle \mathbf{b}_i, v_i \rangle \in B_{\text{su}} - V_{\infty}$  **do**
  - (15)          $c(\mathbf{b}_i) \leftarrow \min_{s \in S} \frac{b(s)}{b_i(s)}$
  - (16)      $f(\mathbf{b}_i) \leftarrow v_i - \sum_{\mathbf{b}' \in V_{\infty}} v(\mathbf{b}') b_i(s)$
  - (17)     **end for**
  - (18)      $v \leftarrow v_b^0 + \min_i c(\mathbf{b}_i) f(\mathbf{b}_i)$ 并将点值对 $\langle \mathbf{b}, v \rangle$ 加入上界集合 $\bar{V}$
  - (19) **end for**
- 

最后将 $B_{\text{su}}$ 中每个信念点选出的向量获得 $\Gamma_{t+1} =$

$$\bigcup_{\mathbf{b} \in B_{\text{su}}} \text{up}(\mathbf{b})。$$

基于 $B_{\text{su}}$ 更新值函数向量集 $\Gamma_{t+1}$ 算法示于表2。

更新 $B_{\text{su}}$ 的复杂度近似为 $O(|B_{\text{su}}| \|A\| |Z|)$ , 基于 $B_{\text{su}}$ 而更新值函数向量集合 $\Gamma_t$ 过程的复杂度近似为 $O(|S|^2 |A| |Z| |B_{\text{su}}|^2)$ , 所以整个算法复杂度近似为 $O(|B_{\text{su}}| \|A\| |Z| + |S|^2 |A| |Z| |B_{\text{su}}|^2)$ 。

表2 算法2: 基于 $B_{\text{su}}$ 更新值函数向量集 $\Gamma_{t+1}$

- 
- (1) **for all**  $\mathbf{b} \in B_{\text{su}}$  **do**
  - (2)     向量集合 $\Gamma_{t+1, \mathbf{b}} \leftarrow \emptyset$
  - (3)     **for all**  $a \in A$  **do**
  - (4)         向量 $\Gamma_{t+1, \mathbf{b}}^a \leftarrow 0$
  - (5)         **for all**  $z \in Z$  **do**
  - (6)             用式(18)计算 $\Gamma_{t+1}^{a,z}$
  - (7)              $\Gamma_{t+1, \mathbf{b}}^a \leftarrow \arg \max_{\alpha \in \Gamma_{t+1}^{a,z}} \mathbf{b} \cdot \alpha + \Gamma_1^a$
  - (8)         **end for**
  - (9)     将向量 $\Gamma_{t+1, \mathbf{b}}^a$ 加入集合 $\Gamma_{t+1, \mathbf{b}}$ 中
  - (10)    **end for**
  - (11)    将 $\Gamma_{t+1, \mathbf{b}}$ 中与 $\mathbf{b}$ 相乘最大的向量加入 $\Gamma_{t+1}$
  - (12) **end for**
-

## 5 仿真结果与分析

仿真中随机生成两个物理网络拓扑分别作为DU池和CU池的基础设施网络, DU池和CU池的节点数为16和23, 并设前传网络的带宽资源总量为260 Mb/s。任意节点的计算资源和任意链路的带宽资源随机取值, 计算资源由CPU个数定义, 其CPU的个数取值范围为(1, 4), 带宽资源的取值范围为(10 Mb/s, 20 Mb/s)。仿真场景中有3个不同接入网切片, 各切片的SFC数量比例为5:3:2, 且系统的SFC总数量取值范围为[10, 70], 且最大队列长度为10 Mb。同时, 将仿真的迭代次数设置为200次。

图3是3种SFC部署方案随着SFC数目的吞吐量情况。从图中可以看出, 在相同SFC数目情况下, POMDP的吞吐量是最大的, 切片规模较小时增加较快, 较大规模时吞吐量的增加减慢。GA模型的吞吐量小于POMDP, 在SFC的数目为70时吞吐量急剧下降, MA方案的吞吐量小于GA, 且随SFC数目增大时吞吐量增加缓慢, 甚至还会下降。对比可知, 基于POMDP的SFC部署方法可以提高吞吐量。

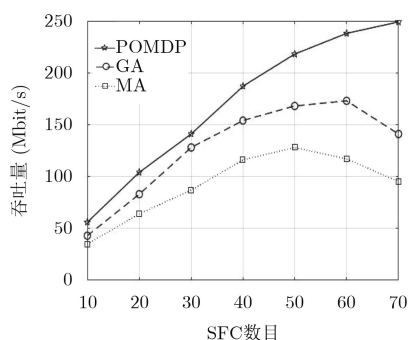


图3 3种SFC部署方案的吞吐量

图4给出3种不同方案SFC资源分配算法的资源利用率情况。从柱形图可以得到, POMDP方案在计算资源和带宽资源的利用率都是最高的, 验证了该部署算法可以充分利用接入网的网络资源。GA方案资源利用率低于POMDP, 说明基于核心网的切片部署算法无法很好地支持接入网切片部署。此外, 静态部署的MA方案在3种算法中, 其资源利用率最低。

图5评估了解决POMDP模型的3种算法的性能。在节点规模为(16,23)和SFC总数为70的场景下, 由图可知, 依赖密度的PBVI(Point-Based Value Iteration)算法的收敛速度是最慢的, 且所得解也是最差的。只依赖值函数的HSVI(Heuristic

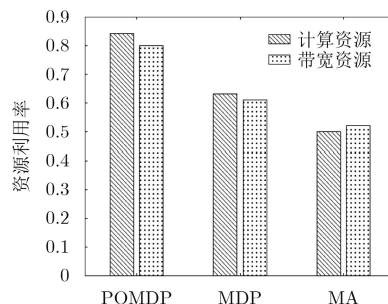


图4 3种SFC资源分配算法方案的资源利用率

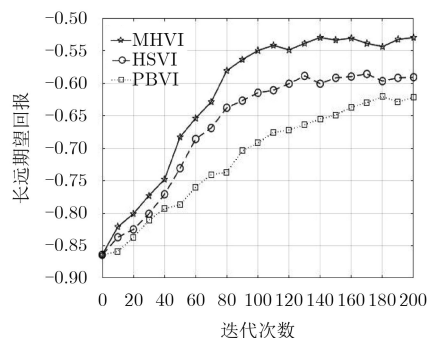


图5 3种POMDP求解算法对比

Search Value Iteration)算法其收敛速度和收敛值明显优于PBVI算法。MHVI算法在HSVI算法基础上改进而成的, 在收敛速度上比HSVI算法稍快, 且寻找最优解的能力优势明显, 验证了MHVI算法性能的优越性。

图6给出了接入网VNF部署方式的选择频率和切片服务数据规模的关系。由图可知, 切片1服务数据规模最大, DU里完成的功能偏多从而减小对前传网络的带宽资源的消耗, 所以接入网VNF部署方式集中在3, 4和5。切片2服务数据规模居中, 对前传网络资源依赖性减弱, 所以优先选择选项5。切片3的数据规模最小, 对前传资源依赖性最弱, 在DU池比CU池的网络资源少的情况下, 尽量在CU池里实现功能, 所以切片3选择选项6的频率最高。

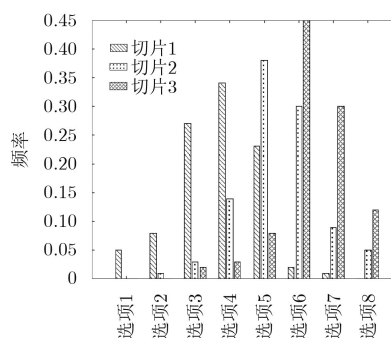


图6 3个切片的接入网VNF部署方式统计图

## 6 结束语

本文考虑在5G接入网CU/DU架构场景下, 针对接入网切片的时延问题, 提出了基于POMDP拓扑感知的接入网切片的SFC部署方案。该方案中引入心跳包检测机制部分感知网络拓扑, 建立一个优化部署SFC的POMDP模型, 降低切片时延。并采用基于点的混合启发式算法求解。仿真结果表明, 这种部署方案能提升网络的吞吐量和资源利用率。

### 参考文献

- [1] SHARMA S, MILLER R, and FRANCINI A. A cloud-native approach to 5G network slicing[J]. *IEEE Communications Magazine*, 2017, 55(8): 120–127. doi: [10.1109/MCOM.2017.1600942](https://doi.org/10.1109/MCOM.2017.1600942).
  - [2] ZHANG Haijun, LIU Na, and CHU Xiaoli. Network slicing based 5G and future mobile networks: Mobility, resource management, and challenge[J]. *IEEE Communications Magazine*, 2017, 55(8): 138–145. doi: [10.1109/MCOM.2017.1600940](https://doi.org/10.1109/MCOM.2017.1600940).
  - [3] KATSALIS K, NIKAEIN N, and SCHILLER E. Network slices toward 5G communications: Slicing the LTE network[J]. *IEEE Communications Magazine*, 2017, 55(8): 146–154. doi: [10.1109/MCOM.2017.1600936](https://doi.org/10.1109/MCOM.2017.1600936).
  - [4] FOUKAS X, PATOUNAS G, and ELMOKASHFI A. Network slicing in 5G: Survey and challenges[J]. *IEEE Communications Magazine*, 2017, 55(5): 94–100. doi: [10.1109/MCOM.2017.1600951](https://doi.org/10.1109/MCOM.2017.1600951).
  - [5] LI Xin and SAMAKA M. Network slicing for 5G: Challenges and opportunities[J]. *IEEE Internet Computing*, 2017, 21(5): 20–27. doi: [10.1109/MIC.2017.3481355](https://doi.org/10.1109/MIC.2017.3481355).
  - [6] MIJUMBI R, SERRAT J, and GORRICO J L. Network function virtualization: state-of-the-art and research challenges[J]. *IEEE Communications Surveys Tutorials*, 2017, 18(1): 236–262. doi: [10.1109/COMST.2015.2477041](https://doi.org/10.1109/COMST.2015.2477041).
  - [7] GIL J H and BOTERO J F. Resource allocation in NFV: A comprehensive survey[J]. *IEEE Transactions on Network and Service Management*, 2016, 13(3): 518–532. doi: [10.1109/TNSM.2016.2598420](https://doi.org/10.1109/TNSM.2016.2598420).
  - [8] HUANG Huawei and SONG Guo. Service chaining for hybrid network function[J]. *IEEE Transactions on Cloud Computing*, 2017. doi: [10.1109/TCC.2017.2721401](https://doi.org/10.1109/TCC.2017.2721401).
  - [9] QU Long, ASSI C, and SHABAN K. Delay-aware scheduling and resource optimization with network function virtualization[J]. *IEEE Transactions on Communications*, 2016, 64(9): 3746–3758. doi: [10.1109/TCOMM.2016.2580150](https://doi.org/10.1109/TCOMM.2016.2580150).
  - [10] MAHMOOD A M, AL-YASIRI A, and ALANI O Y K. A new processing approach for reducing computational complexity in cloud-RAN mobile networks[J]. *IEEE Access*, 2018, 6: 6927–6946. doi: [10.1109/ACCESS.2017.2782763](https://doi.org/10.1109/ACCESS.2017.2782763).
  - [11] CHIH I. RAN revolution with NGFI (xhaul) for 5G[J]. *Journal of Lightwave Technology*, 2018, 36(2): 541–550. doi: [10.1109/JLT.2017.2764](https://doi.org/10.1109/JLT.2017.2764).
  - [12] ZHANG Nan, LIU Yafeng, and FARMANBAR H. Network slicing for service-oriented networks under resource constraints[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(11): 2512–2521. doi: [10.1109/JSAC.2017.2760147](https://doi.org/10.1109/JSAC.2017.2760147).
  - [13] HAYASHIBARA N, DEFAGO X, and YARED R. The  $\varphi$  accrual failure detector[C]. *IEEE International Symposium on Reliable Distributed Systems*, Florianopolis, Brazil, 2014: 66–78.
  - [14] 刘峰. 基于部分可观察马尔科夫决策过程的序列规划问题的研究[D]. [博士论文], 南京大学, 2015.  
LIU Feng. A study of sequence planning based on partially observable markov decision process[D]. [Ph.D. dissertation], Nanjing University, 2015.
  - [15] CILDEN E and POLAT F. Toward generalization of automated temporal abstraction to partially observable reinforcement learning[J]. *IEEE Transactions on Cybernetics*, 2017, 45(8): 1414–1425. doi: [10.1109/TCYB.2014.2352038](https://doi.org/10.1109/TCYB.2014.2352038).
  - [16] ZHENG Qiang, ZHENG Kan, ZHANG Haijun, et al. Delay-optimal virtualized radio resource scheduling in software-defined vehicular networks via stochastic learning[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(10): 7857–7867. doi: [10.1109/TVT.2016.2538461](https://doi.org/10.1109/TVT.2016.2538461).
- 陈前斌: 男, 1967年生, 教授, 博士生导师, 主要研究方向为个人通信、多媒体信息处理与传输、下一代移动通信网络。  
杨友超: 男, 1993年生, 硕士生, 研究方向为网络虚拟化和切片资源分配。  
周 钰: 男, 1993年生, 硕士生, 研究方向为切片资源分配和深度学习。  
赵国繁: 女, 1993年生, 硕士生, 研究方向为5G网络切片中的资源分配、可靠性。  
唐 伦: 男, 1973年生, 教授, 博士生导师, 主要研究方向为新一代无线通信网络、异构蜂窝网络。