文章编号:1671-4598(2025)04-0299-07 DOI:10.16526/j.cnki.11-4762/tp.2025.04.040 中图分类号:TP393 文献标识码:A

基于 SMDP 的安全防护基础设施网络切片服务响应模型

王璐茜，葛洪武，朱晓明，贾 哲
（中国电子科技集团公司 第 54 研究所，石家庄 050081）

摘要：在创建各种安全防护基础切片络切片的过程中，安全防护基础设施提供商主要负责响应切片服务请求和分配资源，但其切片资源有限，切片服务存在响应速度慢的问题；因此，针对安全防护基础设施提供商的服务决策过程，提出了一种基于半马尔可夫决策过程的安全防护基础设施网络切片服务响应模型；经过仿真分析，基于 SMDP 的模型在有限的资源条件下，能够提高切片资源利用效率并最大化收益，高效地响应切片服务请求。

关键词：安全防护基础设施；网络切片；半马尔可夫决策过程；安全防护基础设施提供商；服务响应模型

An SMDP Based Slicing Service Response Model for Safeguarding infrastructure Network

WANG Luxi, GE Hongwu, ZHU Xiaoming, JIA Zhe
(The 54th Research Institute of China Eectronics Tchnology Group Corporation, Shijiazhuang 050081, China)

Abstract: During the process of slicing the network for protective basic slices, the security protection infrastructure providers are mainly responsible for responding to slice service requests and allocating resources. However, due to the limited slicing resources they possess, there exists a problem of slow response speed in slice services. Therefore, with respect to the service decision-making process of security protection infrastructure providers, a security protection infrastructure network slice service response model based on the Semi-Markov Decision Process (SMDP) is proposed. Through simulation analysis, it is demonstrated that the model based on SMDP, under the condition of limited resources, is capable of enhancing the utilization efficiency of slicing resources, maximizing profits, and efficiently responding to slice service requests.

Keywords: safeguarding infrastructure; network slicing; semi-Markov decision process; safeguarding infrastructure providers; service response model

0 引言

随着信息安全技术的发展，网络攻击事件频频发生，安全防护基础设施在网络安全防护中的作用越来越显著，安全防护基础设施是指一系列硬件、软件、策略和程序的多元网络，包含应用防火墙、入侵检测系统、流量探针、漏扫、蜜罐、网络隔离设备、虚拟专用网络等各种安全元素，旨在保护计算机网络及其传输的数据免受未授权访问、攻击、破坏或篡改等威胁。这些安全防护基础设施是确保信息保密性、完整性和可用性的重要保障，在维护个人隐私、企业资产和国家安全方面具有关键作用。然而，网络攻击场景的多样化，导致不同的安全用户对安全防护基础设施提出了各种各样的定制化服务质量要求，例如，对于企业网络，需要保护敏感数据和关键业务系统，防火墙作为第一道防线，控制进出网络的流量，入侵检测系统，实时监控和防御网络攻击，虚拟专用网络（VPN，virtual private network），为远程员工提供安全的网络接入，数据丢失防护软件，防止敏感数据外泄；对于工业控制系统（ICS，industrial control system），需要保护关键基础设施，通过网

收稿日期:2025-03-25； **修回日期**:2025-04-07。

作者简介:王璐茜(2000-),女,硕士研究生。

引用格式:王璐茜,葛洪武,朱晓明,等.基于 SMDP 的安全防护基础设施网络切片服务响应模型[J].计算机测量与控制,2025,33(4):299-305.

络隔离设备，将工业控制网络与企业网络隔离，减少攻击面，使用流量探针实时监控工业控制网络中的异常活动，通过漏扫，定期更新和修补系统漏洞。针对以上需求，网络切片技术可以依托一种通用的安全防护基础设施物理网络，构建多个相互隔离的专用虚拟化逻辑网络[1]，从而有效契合不同安全客户对网络能力的差异化需求[2]。

通过网络切片，不同的网络安全服务可以由不同的安全防护基础设施网络切片实例提供，每个实例切片由一组虚拟网络功能（VNF，virtualized network function）组成，这些功能运行在共享相同的安全防护基础设施的虚拟资源单元（VRU，virtual resource unit）上，这种方式可以简化网络的部署、运营和管理，并显著地提高服务承载的效率[3]。然而，由于多元化安全用户的需求需要通过相同的基础设施来满足，安全防护基础设施的资源有限且成本高昂，因此，安全防护基础设施提供商响应所有的切片服务请求并为其分配资源的同时，如何提高资源效率并最大化收益成为一个亟待解决的问题。

软件定义网络（SDN，software-defined networking）与软件虚拟化（NFV，network functions virtualization）是网络切片技术的基础[4]，通过虚拟化网络资源和网络功能，进而提高网络资源利用率和灵活性，网络切片编排和管理技术（NSO，network slice orchestration and management）负责切片的创建，监控和管理，确保切片服务的可靠性和稳定性[4]。

软件定义网络技术是一种网络管理方法。建立在将网络基础设施的控制面与转发面分离的基础上，将自动化和编程应用于控制面，使管理员具有动态调整全网流量的能力。通过将网络控制集中到 SDN 控制器上，使整个网络在逻辑上体现为单一的网络设备，通过可编程配置实现自动化配置、控制、保护和资源调整。在软件定义网络架构体系中，控制平面借助控制－转发通信接口，对网络设备实施集中式管控，控制器与网络设备间交互的控制信令所产生的流量，与终端间通信引发的数据流量相互独立[5]。网络设备接收来自控制器的控制信令后，构建相应的转发表，并依据该转发表对数据流量的处理方式做出决策[5]。这一机制的应用，规避了网络设备在数据转发过程中对复杂分布式网络协议的依赖，简化了传统网络数据转发流程[5]。

网络功能虚拟化（NFV）是一种关于网络架构的概念[6]。通常企业使用的服务器由硬件厂商统一生产，通过后续部署不同的操作系统以及安装不同的软件来实现丰富各异的功能[7]。然而在安全防护基础设施中，许多常见的网络设备并没有采用这种模式，防火墙、入侵检测系统、流量探针、漏扫、蜜罐、网络隔离设备、路由器、交换机等软硬件设备均有自己独立的硬件和软件系统[6]。网络功能虚拟化在架构设计上汲取了传统服务器架构理念，将路由器、交换机、防火墙、负载均衡这些不同的网络功能封装成独立的可迁移的模块化软件，即虚拟网络功能，这些虚拟网络功能可以根据实际需求灵活部署、迁移和扩展[7]。通过在硬件设备上运行不同的模块化软件，能够突破传统架构的局限在单一硬件设备上承载多样化的网络功能[7]，提升了网络资源的利用率和灵活性。

网络切片编排和管理技术（NSO）是切片网络中的核心使能技术，旨在通过自动化、智能化的方式，实现对网络切片的全生命周期动态创建、部署、优化和运维，从而有效契合不同业务场景不同安全客户在网络性能方面的差异化需求。

目前，业界关于网络切片的资源管理与分配的研究较为广泛，文献［8］在云无线接入网（C-RAN，cloud radio access network）场景下，进行了端到端切片资源分配的相关研究，该研究通过分析两个不同层次的紧密耦合问题来设计资源分配方案，并开发了一种快速算法来分配前传容量和云计算资源。在文献［9］中，提出了一种基于破产博弈的算法来为网络切片分配资源。在该博弈中，云和切片分别被建模为破产公司和债权人[4]，并采用 Shapley 值来获得稳定的解决方案。然而，这些研究并未考虑切片服务的请求和结果。文献［10］和文献［11］提出了虚拟网络功能（VNF）放置和切片请求映射的维度，该研究的目标仅是提高资源分配效率，且默认接受所有切片服务请求。此外，文献［12］中设计了一种商业模式，探讨了利润最大化和资源效率之间的关系。然而，从安全防护基础设施切片动态资源分配过程的角度上看，针对切片服务的随机开始、结束以及系统资源容量的变化的场景，应当兼顾考虑系统收益、资源效率和切片服务响应，但业界在此方面研究较少。

马尔科夫分析（MA，markov analysis）可以被用来解决上述问题，是一种适用于系统状态转移建模并借助该模型计算系统到达各种状态的概率的分析技术[13]。MA 是进行复杂系统建模的有力工具，广泛应用于涉及时序、顺序、修理、冗余和容错机制的系统场景[13]。在实际应用中，通常运用 MA 来绘制系统状态转移图，并对其进行深入分析，从而确定系统是如何到达非预期的状态以及计算相应的概率[13]。MA 可用于系统性能、可信性、可用性、可靠性以及安全性的建模，描述系统的故障状态和降级运行状态，其中降级状态是指系统或者部分故障，或者只能执行部分功能而其他功能则无法

完成[14]。

马尔科夫决策过程（MDP，markov decision process）在解决网络安全防护基础设施网络切片服务响应问题时主要用于建模决策模型。它将该问题建模为一个动态系统，该系统的状态是随机的，需要做出决定，其代价由决策决定[14]。

但是在网络安全防护基础设施网络切片服务响应的决策问题中，决策阶段之间的时间不是恒定的，而是随机的[14]，传统的马尔科夫决策过程不能很好地进行建模。而半马尔可夫决策过程（SMDP，semi-markov decision process）是马尔科夫决策过程的一种扩展，通常用于对随机控制问题进行建模，相比于马尔科夫决策过程，半马尔科夫决策过程的每个状态都具有一定的逗留时间，并且逗留时间是一个通用的连续随机变量[14]，更适合用于对网络安全防护基础设施网络切片服务响应问题进行建模。

因此，本文提出了一种基于半马尔可夫决策过程（SMDP）的切片资源管理方法，构建了一个全面的安全防护基础设施网络切片服务响应模型，在保证服务响应的情况下，通过提高切片资源利用，达到系统收益的最大化。

1　网络切片服务响应模型

在本节中，提出了最优服务决策的切片服务响应模型。首先，介绍了安全防护基础设施网络切片服务响应的整体过程，然后，描述了模型的系统状态以及每个状态下的动作，最后，评估了安全防护基础设施提供商（SinP，safeguarding infrastructure providers）关注的系统收益。

1.1　模型描述

如图 1 所示，本研究考虑一个由多个虚拟网络功能（VNF）组成的安全防护基础设施网络切片，不同的切片代表不同的服务类型。VNF 部署在安全防护基础设施网络提供的虚拟资源单元（VRU）中，VNF 的数量是可变的，取决于 VRU 的容量和服务类型。当安全用户发送服务请求时，系统会评估预期系统收益、预期系统开销以及在计算期间占用 VRU 的利用率，以决定是否接受或拒绝该服务请求。如果决定接受，安全防护基础设施提供商（SinP）将为此服务分配 VRU。对于该模型，假设基础设施网络中有 K 个 VRU，一个切片服务请求占用 c 个 VRU，其中 $c \in \{1,2,3\cdots C\}, C \leqslant K$。此外，假设切片服务的到达率服从参数为 λ 的泊松分布，服务时间服从参数为 μ 的指数分布。

1.2　系统状态

已分配 c 个虚拟资源单元（VRU）的服务数量记为 n_c，系统中占用的 VRU 总数为 $\sum_{c=1}^{C}(n_c * c)$，其中是可

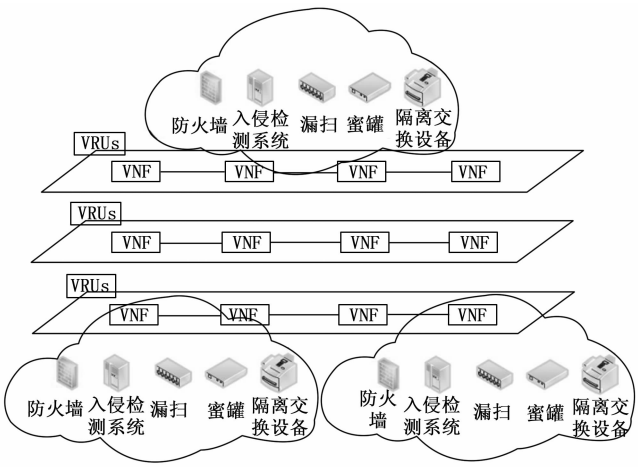

图 1　安全防护基础设施网络切片部署图

以分配给一个服务请求的最大 VRU 数量，R 表示新服务请求的到达，而当切片服务结束时，占用的 VRU 将被释放，记为 D_c。系统中共有两种类型的事件，可一起表示为 $e = \{R, D_1, D_2, \cdots, D_C\}$，系统状态可以通过当前具有不同数量 VRU 的服务以及系统中的事件来表征，这些事件可能是切片服务请求到达或结束。

$$S = \{s \mid s = [n_1 n_2, \cdots, n_C, e]\} \tag{1}$$

1.3　系统收益

当请求切片服务到达时，可以从动作集 A 中选择两种动作：接受并分配 c 个虚拟资源单元（VRU）或拒绝，可表示为：

$$A(s) = \begin{cases} 0,1,2\cdots C, & e \in R \\ -1, & e \in D \end{cases} \tag{2}$$

在网络切片服务响应模型中，系统收益是基于收入和成本的即时利润，记为 $r(s,a)$，其计算公式为：

$$r(s,a) = w(s,a) - g(s,a) \tag{3}$$

其中：$S = \{s \mid s = < n_1, n_2, \cdots, n_C, e >\}$，$w(s,a)$ 是当状态 s 中发生事件 e 时，安全防护基础设施提供商（SinP）通过做出决策或行动所获得的总收入。$g(s,a)$ 是预期的系统成本，其计算公式为：

$$g(s,a) = \tau(s,a) o(s,a) \tag{4}$$

其中：$\tau(s,a)$ 是做出决策时的预期服务时间；$o(s,a)$ 是服务时间的成本率，它由占用的虚拟资源单元（VRU）数量决定：

$$o(s,a) = \sum_{c=1}^{C}(n_c * c) \tag{5}$$

总收入 $w(s,a)$ 的计算公式为：

$$w(s,a) = \begin{cases} 0, & A(s) = -1, \quad e \in D \\ -1, & A(s) = 0, \quad e \in R \\ E - \beta/c\mu, & A(s) = c, \quad e \in R \end{cases} \tag{6}$$

当网络切片服务完成时，安全防护基础设施提供商（SinP）是不会获得收益的，因此 $A(s) = -1$，$w(s,a)$

$=0$。当新的服务请求到达时，将获得收益 E，同时该服务将占用 c 个虚拟资源单元（VRU），这涉及的资源占用费用为 $\frac{\beta}{c\mu}$，其中 β 表示单位时间的价格。当 $e\in R$ 且 $A(s)=0$ 时，表示服务请求转换费用 U[15]。

2 半马尔可夫决策过程（SMDP）建模

在本节中，我们基于半马尔可夫决策过程（SMDP）建模了状态转移概率，并最大化安全防护基础设施提供商（SinP）的系统收益。随后，提出了一种强化学习算法来求解最大收益。最后，我们通过优化切片服务的阻塞概率，提高了网络切片服务请求的响应速度。

2.1 状态转移概率

$q(j\mid s,a)$ 表示在采取动作 A 后，从状态 s 转移到状态 j 的状态转移概率。对于状态 $s=(n_1,n_2,\cdots,n_c,\cdots,n_C,R)$，$q(j\mid s,a)$ 可以通过以下方式获得：

$$q(j\mid s,a)=\begin{cases}\dfrac{\lambda}{\gamma(s,a)} & j=(n_1,n_2,\cdots,n_C,R)\\ \dfrac{n_c c\mu}{\gamma(s,a)} & j=(n_1,\cdots,n_c-1,\cdots,n_C,D_c)\\ \dfrac{(n_c+1)c\mu}{\gamma(s,a)} & j=(n_1,\cdots,n_c,\cdots,n_C,D_c)\\ \dfrac{n_m c\mu}{\gamma(s,a)} & j=(n_1,\cdots,n_m-1,n_c+1,\cdots,n_C,D_c)\\ \dfrac{\lambda}{\gamma(s,a)} & j=(n_1,\cdots,n_c+1,\cdots,n_C,R)\end{cases}\tag{7}$$

其中：$m\neq c$，γ 表示切片服务到达率，μ 表示切片服务率开率，其他变量的定义与 1.2 节和 1.3 节相同。

对于状态 $s=(n_1n_2,\cdots,n_c,\cdots,n_C,D_c)$，当切片服务完成时，除了 $A(s)=-1$ 之外没有其他动作；因此，$q(j\mid s,a)$ 可以通过以下方式获得：

$$q(j\mid s,a)=\begin{cases}\dfrac{\lambda}{\gamma(s,a)} & j=[n_1,n_2,\cdots,n_C,R]\\ \dfrac{n_c c\mu}{\gamma(s,a)} & j=[n_1,\cdots,n_c-1,\cdots,n_C,D_c]\end{cases}\tag{8}$$

其中：$c\in\{1,2,\cdots,C\}$。

2.2 最大系统收益

给定当前状态 s 和所选决策 A，我们将从当前时刻到下一时刻的时间间隔记为 $\tau(s,a)$。因此，对于给定的 s 和 A，事件的平均速率记为 $\gamma(s,a)$，它是整个系统中所有事件速率的总和，计算公式为：

$$\gamma(s,a)=\tau(s,a)^{-1}=\begin{cases}\lambda+\sum_{c=1}^{C}n_c cu & e\in R,\quad A(S)=0\\ & e\in\{D_1,D_2\cdots,D_C\}\\ \lambda+\sum_{c=1}^{C}n_c cu+cu & e\in R,\quad A(S)=c\end{cases}\tag{9}$$

如上所述，两个连续决策时刻之间的时间间隔服从指数分布，因此基于折扣奖励模型计算系统预期如下：

$$r(s,a)=w(s,a)-o(s,a)E_s^a\int_0^\tau e^{-\alpha t}\,\mathrm{d}t=w(s,a)-o(s,a)E_s^a\left\{\frac{[1-e^{-\alpha\tau}]}{\alpha}\right\}=w(s,a)-\frac{o(s,a)}{\alpha+\gamma(s,a)}\tag{10}$$

其中：α 是连续时间折扣因子。进一步地，可以获得最大系统收益，即贝尔曼方程[16]：

$$v(s)=\max_A\left[r(s,a)+\varepsilon\sum_{j\in S}q(j\mid s,a)v(j)\right]\tag{11}$$

其中：$\varepsilon=\gamma(s,a)/[\gamma(s,a)+\alpha]$。

通过统一化后推导出 $v(s)$ 的最优系统收益方程为：

$$\tilde{v}(s)=\max_A\left[\tilde{r}(s,a)+\tilde{\varepsilon}\sum_{j\in S}\tilde{q}(j\mid s,a)\tilde{v}(j)\right]\tag{12}$$

其中：折扣奖励函数和状态转移概率分别表示为：

$$\tilde{r}(s,a)=r(s,a)[1+\alpha\tau(s,a)]/[(\alpha+\omega)\tau(s,a)]\tag{13}$$

$$\tilde{q}(j\mid s,a)=\begin{cases}1-\dfrac{[1-q(s\mid s,a)]}{\tau(s,a)\omega} & j=s\\ \dfrac{q(s\mid s,a)}{\tau(s,a)\omega} & j\neq s\end{cases}\tag{14}$$

2.3 强化学习算法

考虑到安全防护基础设施网络切片服务响应系统的动态特性，当前状态下的动作选择可能会直接导致下一状态的剧烈变化，并对长期系统预期收益产生重大影响，从长远来看，采取仅最大化当前状态奖励的动作是不明智的。因此，本文提出了一种强化学习算法，以最大化长期系统预期收益，如算法 1 所示，首先，初始化 $V(S)=0$，对于每个状态 s；计算每个状态的 $V(S)$ 值，并选择最优动作以最大化局部收益。然后将计算得到的收益与之前的值进行比较，直到其收敛为止，退出循环。最后，输出最大长期系统预期收益 $V(S)$[17]。

算法 1：

Input：$q(j\mid s,a)$，$r(s,a)$，$\boldsymbol{A}$，S

Output：Maximum $V(s)$

Initialize $V(S)=0$ for all state $s\in S$

Set $\Delta\geqslant 0$，where Δ is a small number.

for $s\in S$ do

  $v=V(S)$

  $V(S)=\max_{A\in A}\{r(s,a)+\in\sum_{s\in j}q(j\mid s,a)V(j)\}$

  if $|v-V(s)|\leqslant\Delta$ then

    Output Maximum $V(s)$

    break

  End if

 End for

return $V(s)$

2.4　网络切片服务请求阻塞概率

以上已经介绍了系统模型和求解算法。接下来，分析一个重要的性能指标——阻塞概率。当切片服务请求被拒绝时，系统会产生一定的开销，安全用户将无法使用切片服务，这是我们不希望看到的情况。因此，本研究基于提出的半马尔可夫决策过程（SMDP）推导阻塞概率的公式。假设 $\pi_{(n_1 n_2,\cdots,n_c,\cdots,n_C,e)}$ 表示系统中状态 $s=(n_1 n_2,\cdots,n_c,\cdots,n_C,e)$ 的稳态概率 $s=(n_1 n_2,\cdots,n_c,\cdots,n_C,e)$，其中 $\bar{n}=(n_1 n_2,\cdots,n_C)$。需要考虑以下 3 种情况：当一个服务完成时，新的请求到达；拒绝一个新的服务请求；接受一个新的服务请求。

$\pi_{(\bar{n},e)}$ 将被分为 $\pi_{(\bar{n},R)}$ 和 $\pi_{(\bar{n},D)}$，基于公式（7）（8）中推导的转移概率，可以得到 $\pi_{(\bar{n},R)}$ 和 $\pi_{(\bar{n},D)}$。如下所示：

$$\pi_{(\bar{n},R)}=\frac{\lambda}{\gamma(s,a)}\rho(\bar{n},R)\pi_{(\bar{n},R)}+\frac{\lambda}{\gamma(s,a)}\sum_{c=1}^{C}\rho(\overline{n-1},R)\pi_{(\overline{n-1},R)}+\frac{\lambda}{\gamma(s,a)}\sum_{c=1}^{C}\pi_{(\bar{n},D)} \tag{15}$$

式中，$\rho(\bar{n},R)$ 和 $\rho(\overline{n-1},R)$ 被定义为：

$$\rho(\bar{n},R)=\begin{cases}1, & A(S)=0\\ 0, & \text{otherwise}\end{cases} \tag{16}$$

$$\rho(\overline{n-1},R)=\begin{cases}1, & A(S)=c;c=\{1,2,\cdots,C\}\\ 0, & \text{otherwise}\end{cases} \tag{17}$$

类似地，稳态概率 $\pi_{(\bar{n},D)}$ 可以表示为：

$$\pi_{(\bar{n},D)}=\frac{(n_c+1)c\mu}{\gamma(s,a)}\rho(\overline{n+1},R)\pi_{(\overline{n+1},R)}+\frac{(n_c+1)c\mu}{\gamma(s,a)}\rho(\bar{n},R)\pi_{(\bar{n},R)}+\frac{(n_m+1)c\mu}{\gamma(s,a)}\sum_{m=1,m\neq c}^{C}\rho(\overline{n\pm1},R)\pi_{(\overline{n\pm1},R)}+\frac{(n_c+1)c\mu}{\gamma(s,a)}\sum_{c=1}^{C}\pi_{(\overline{n+1},D)} \tag{18}$$

其中：$\rho(\bar{n},R)$，$\rho(\overline{n+1},R)$ 和 $\rho(\overline{n\pm1},R)$ 被定义为：

$$\rho(\bar{n},R)=\begin{cases}1,A(S)=c;c=\{1,2,\cdots,C\}\\ 0,\text{otherwise}\end{cases}$$

$$\rho(\overline{n\pm1},R)=\begin{cases}1,A(S)=c;c=\{1,2,\cdots,C\}\\ 0,\text{otherwise}\end{cases}$$

$$\rho(\overline{n+1},R)=\begin{cases}1,A(S)=0;\\ 0,\text{otherwise}\end{cases} \tag{19}$$

所有稳态概率的总和等于 1。

$$\sum_{S}[\pi_{(\bar{n},R)}+\pi_{(\bar{n},D)}]=1 \tag{20}$$

通过求解方程（14）～（19），可以得到每个状态的稳态概率。阻塞概率由 P_0 表示，它是被拒绝的新切片服务的总概率与到达的新切片服务的总概率之比，当 $A(S)=0$ 时，所有稳态概率的总和表示被拒绝服务的概率，可以得出：

$$P_0=\frac{\sum_{A(S)=0}\pi_{(\bar{n},R)}}{\sum_{A}\pi_{(\bar{n},R)}} \tag{21}$$

3　实验与结果分析

在本节中，使用 Matlab 2018b 进行仿真实现，固定参数选择如表 1 所示。为了进一步研究这些参数之间的关系，我们改变了一些参数，例如切片服务到达率和结束率。假设安全防护基础设施提供商（SinP）包含 4 个虚拟资源单元（VRU），并且具有虚拟网络功能（VNF）的最大服务将需要两个 VRU，即 $C=2$。同时，折扣因子设置为 $\alpha=0.1$，以确保奖励的收敛性。在此基础上，进行了大量实验，得到了以下结果。

表 1　固定参数表

基础收益	转换费用	时间折扣因子	单位时间价格
E	U	α	β
6	2	0.1	1

如图 2 所示，我们将半马尔可夫决策过程（SMDP）的系统预期收益与启发式方法进行了比较，其中，将虚拟资源单元（VRU）的总和设置为 $K=4$，服务的结束率 $\mu=2$；通过对比可以看出，当切片服务到达率 λ 较低时，贪婪算法的奖励高于模拟退火（SA）算法，这是因为服务到达的数量较少，系统有足够的资源进行分配，因此贪婪分配的奖励较高，但是切片服务到达率的增加，本文提出的模型优于启发式方法，并且最大系统预期收益随着 λ 的增加而增加，表现一直优于其他两种算法。

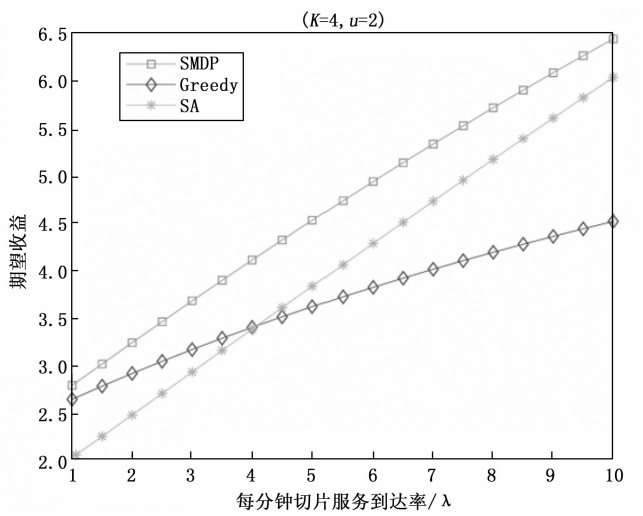

图 2　不同切片服务到达率的期望收益

此外，我们将虚拟资源单元（VRU）的总数表示

为一个变量值，通过每次改变 VRU 数量 k 值，可以获得更多的状态。选择一个固定状态 $(1,0,R)$，并在每个 K 下观察其最大系统预期收益，如图 3 所示，其中 $\lambda=7$ 时 $\mu=2$ 。当 VRU 的总数较低时，奖励也会非常低，系统无法提供更多的虚拟资源。一些网络切片无法部署在 VRU 中。随着 k 的增加，更多的服务被接受，奖励也随之增加。从图 4 可以看出，本文提出的模型显著优于贪婪策略。原因是，对于贪婪策略，当服务请求到达时，会立即分配最大数量的 VRU，因此在可用 VRU 不足时，存在拒绝下一个切片服务请求的风险。

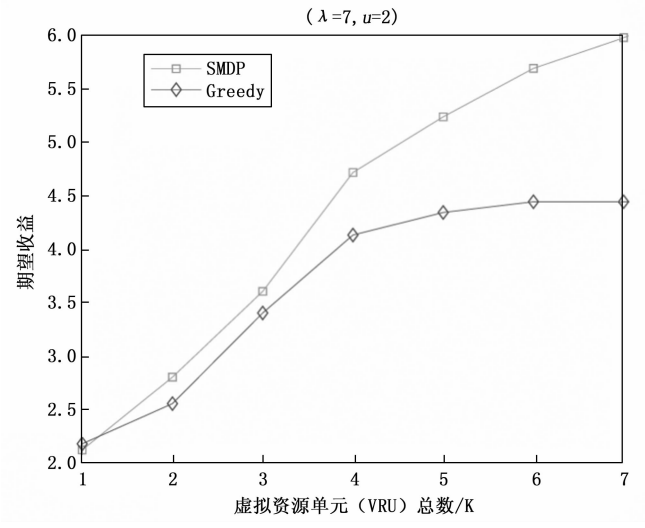

图 3 不同 VRU 数量的期望收益

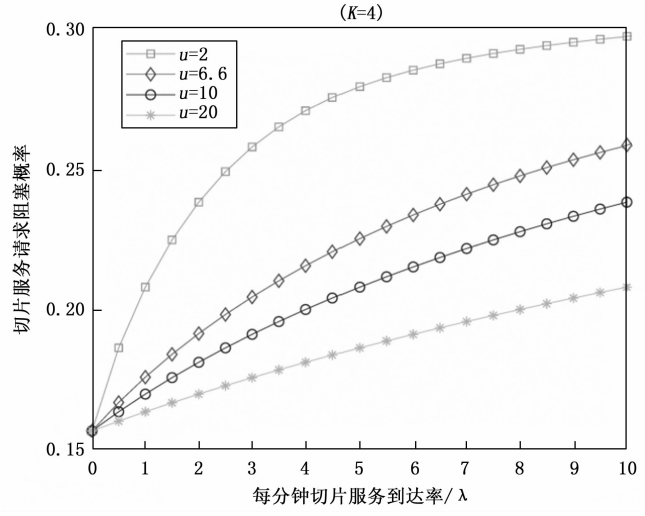

图 4 不同切片服务到达率下的阻塞概率

由图 4 可以看出，由于系统资源有限，当新服务到达率增加时，安全防护基础设施提供商（SinP）没有足够的资源提供相应的服务，因此更多的服务可能会被拒绝，阻塞概率也随之升高。然而，服务处理时间会在切片服务结束率增加的同时有所缩短，这意味着系统将能够在单位时间内处理更多的切片服务[18]，阻塞概率会随之降低。针对本文提出的基于 SMDP 的安全防护基础设施网络切片服务响应模型，随着切片服务到达率的增加，服务请求阻塞概率也在增加，但会趋近于一定的值，不会随着服务到达率的增加而无限增加。因此，保证了切片服务响应速度，保障了安全用户的不同服务需求。

4 结束语

本文提出了一种基于半马尔可夫决策过程（SMDP）的切片服务响应模型，用于安全防护基础设施网络切片中的资源动态分配，并结合强化学习算法，来求解最大系统预期收益。同时，该模型通过优化切片服务的阻塞概率，在提高资源效率利用的前提下，保证了网络切片服务请求的响应速度，以满足更多安全用户的需求。

参考文献：

[1] 李红祎，赵一荣，李金艳，等. 基于能力开放的 5G 网络切片管理研究 [J]. 电子技术应用，2020，46 (1)：1－5.

[2] 魏风生. 网络切片智能配置与资源调度研究 [D]. 成都：电子科技大学，2022.

[3] BEGA D，GRAMAGLIA M，BANCHS A，et al. Optimising 5G infrastructure markets：the business of network slicing [C] //IEEE INFOCOM 2017-IEEE Conference on Computer Communications，2017：910－918.

[4] 张 健. 面向 5G 网络的高效节能资源分配算法研究 [D]. 北京：北京邮电大学，2020.

[5] 陈家宣. 基于容器的分布式 SDN 网络控制器技术研究 [D]. 北京：北京邮电大学，2021.

[6] 智次方研究院. 2024 年中国 5G 产业全景图谱报告 [EB/OL]. [2023－12－23]. https：//max. book118. com/html/2023/1220/6133015113010023. shtm.

[7] 刘国旭. 数据传输方法及装置、电子设备、存储介质 [P]. 中国：202211030602. 2，2022－08－25.

[8] NGUYEN HA V，LE L B. End-to-end network slicing in virtualized OFDMA-based cloud radio access networks [J]. IEEE Access，2017，5：18675－18691.

[9] JIA Y，TIAN H. Bankruptcy game-based resource allocation algorithm for 5G cloud-RAN slicing [C] //2018 IEEE Wireless Communications and Networking Conference (WCNC)，2018.

[10] LIN T C，ZHOU Z L，TORNATORE M，et al. Optimal network function virtualization realizing end-to-end requests [C] //2015 IEEE Global Communications Confer-

ence (GLOBECOM)，2015：3108－3113.

[11] COHEN R，LEWIN EYTAN L，NAOR J S，et al. Near optimal placement of virtual network functions [C] // IEEE Conference on Computer Communications，2015：1346－1354.

[12] WANG G，FENG G，TAN W，et al. Resource allocation for network slices in 5G with network resource pricing [C] //2017 IEEE Global Communications Conference，2017：1－6.

[13] 周翔宇. 面向自主船舶的危险分析方法研究 [D]. 大连：大连海事大学，2020.

[14] 杨洪勇. 用于多智能体编队动态路径规划的方法和存储介质 [P]. 中国：202110985514.7，2021－11－23.

[15] LIANG H B，et al. An SMDP-Based Service Model for Interdomain Resource Allocation in Mobile Cloud Networks [J]. IEEE Transactions on Vehicular Technology，2012，61 (5)：2222－2232.

[16] PUTERMAN M L. Markov decision processes：discrete stochastic dynamic programming [M]. New York：John Wiley & Sons，2014.

[17] SUTTON R S，BARTO A G. Reinforcement learning：an introduction [M]. MIT Press，1998.

[18] 周新力. 基于马氏决策过程的多域网络切片资源管理 [D]. 北京：北京邮电大学，2020.

(上接第 298 页)

[4] 穆旭彤，程　珂，宋安霄，等. 抗拜占庭攻击的隐私保护联邦学习 [J]. 计算机学报，2024，47 (4)：842－861.

[5] ZHANG Z，LIU Q，HUANG Z，et al. Model inversion attacks against graph neural networks [J]. IEEE Transactions on Knowledge and Data Engineering，2022，35 (9)：8729－8741.

[6] MCMAHAN B，MOORE E，RAMAGE D，et al. Communication efficient learning of deep networks from decentralized data [C] //Artificial Intelligence and Statistics. PMLR，2017：1273－1282.

[7] 宋凌云，刘至臻，张　炀，等. 基于异构图中多层次图结构的级联图卷积网络 [J/OL]. 软件学报，1－18 [2024－08－21].

[8] CHU T，YANG M，LAOUTARIS N，et al. PriPrune：quantifying and preserving privacy in pruned federated learning [J]. ACM Transactions on Modeling and Performance Evaluation of Computing Systems，2024：2376－3639.

[9] LI T，SAHU A K，TALWALKAR A，et al. Federated learning：Challenges，methods，and future directions [J]. IEEE signal Processing Magazine，2020，37 (3)：50－60.

[10] GAFNI T，SHLEZINGER N，COHEN K，et al. Federated learning：a signal processing perspective [J]. IEEE Signal Processing Magazine，2022，39 (3)：14－41.

[11] YOU Y，WANG H，GU S H，et al. A kind of link tracking log data storage design of monitoring system [J]. Journal of Software，2021，32 (5)：1302－1321.

[12] ZHAO J T，HUANG L S. Research on related technologies of microservice fault diagnosis [J]. Network New Media Technology，2020，9 (1)：57－64.

[13] VERA-RIVERA F H，GAONA C，ASTUDILLO H. Defining and measuring microservice granularity-a literature overview [J]. Peer J. Computer Science，2021，7：e695.

[14] CHEN F，ZHANG L，LIAN X. A systematic gray literature review：the technologies and concerns of microservice application programming interfaces [J]. Software：Practice and Experience，2021，51 (7)：1483－1508.

[15] HE Y H，ZHANG X Y，SUN J. Channel pruning for accelerating very deep neural networks [J]. Proceedings of the 2017 IEEE International Conference on Computer Vision，Venice：IEEE，2017：1398－1406.

[16] LOUIZOS C，WELLING M，KINGMA D P. Learning sparse neural networks through l0 regularization [J]. Arxiv：1712.01312，2017.

[17] ZHU H，ZHANG H，JIN Y. From federated learning to federated neural architecture search：a survey [J]. Complex & Intelligent Systems，2021，7 (2)：639－657.

[18] FENG S，LI B，YU H，et al. Semi-supervised federated heterogeneous transfer learning [J]. Knowledge-Based Systems，2022，252：109384.

[19] CHEN C，LIU Y，MA X，et al. Calfat：calibrated federated adversarial training with label skewness [J]. Advances in Neural Information Processing Systems，2022，35：3569－3581.

[20] ZHOU T，ZHANG J，TSANG D H K. FedFA：federated learning with feature anchors to align features and classifiers for heterogeneous data [J]. IEEE Transactions on Mobile Computing，2023：1－12.

[21] TAN A Z，YU H，CUI L，et al. Towards personalized federated learning [J]. IEEE Transactions on Neural Networks and Learning Systems，2022，34 (12)：9587－9603.