

基于时空关联表征的空天地一体化网络资源精准管控方法

买天乐¹, 姚海鹏², 忻向军¹, 杨景凯², 靳辰朗²

(1. 北京理工大学信息与电子学院, 北京 100081;

2. 北京邮电大学信息与通信工程学院, 北京 100876)

摘要: 空天地一体化网络中, 资源分布不均匀且随时间动态变化。传统基于静态资源表征的优化调度方法仅在单个时间片上进行调度优化, 资源利用率低, 并且在网络变化时需频繁地重新调整策略, 带来较大迁移成本。为解决这一问题, 提出一种空天地一体化网络资源时空关联表征方法。该方法通过在网络的结构邻域和时序邻域两个维度计算节点注意力表征, 实现对空天地一体化网络资源时空关联关系的精准建模。在此基础上, 提出基于时空关联表征的网络资源精准映射方法, 采用置信域策略梯度优化方法, 对算法参数进行自适应学习与优化。

关键词: 空天地一体化网络; 时空注意力模型; 虚拟网络映射; 时变图模型

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-8930.2024014

Spatiotemporal Correlation Representation based Precise Resource Management in Space-Air-Ground Integrated Network

MAI Tianle¹, YAO Haipeng², XIN Xiangjun¹, YANG Jingkai², JIN Chenlang²

1. School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

2. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: In the space-air-ground integrated network, resources are unevenly distributed and dynamically change over time. Traditional optimization scheduling methods based on static resource representation only optimize scheduling within a single time slot, leading to low resource utilization. Moreover, frequent strategy adjustments are required when the network changes, resulting in significant migration costs. To address this issue, this paper proposed a method for spatiotemporal correlation representation of resources in the space-air-ground integrated network. This method computed node attention representations in both structural neighborhoods and chronological neighborhoods, enabled precise modeling of spatiotemporal resource correlations in the space-air-ground integrated network. Based on this, this paper proposed a resource precise embedding method based on spatiotemporal correlation representation, and introduced the trust region policy optimization method to adaptively learn and optimize algorithm parameters.

Keywords: space-air-ground integrated network, spatiotemporal attention model, virtual network embedding, time-varying graph

0 引言

复杂场景(如偏远地区、灾难应急)通信组网需求迫切, 受地形地貌、设施损毁等因素影响, 复杂场景存在基础设施受限(缺失、损毁)、传输路径受阻(易被遮蔽)

等特点, 传统单一制组网方法难以适用, 例如, 光纤网络具有高速大带宽传输能力, 但部署周期长、抗毁性能差; 空中机动网络如无人机基站能够快速灵活部署, 但传输能力有限、链路稳定性差。因此, 亟须构建融合地面网络、空中网络(包括浮空平台和无人机等)以及天基网络(包

收稿日期: 2024-01-31; 修回日期: 2024-06-01

通信作者: 姚海鹏, yaohaipeng@bupt.edu.cn

基金项目: 国家自然科学基金资助项目(No.62325203, No.U22B2033); 国家重点研发计划资助项目(No.2022YFB2902500)

Foundation Items: The National Natural Science Foundation of China (No.62325203, No.U22B2033), National Key Research & Development Program of China (No.2022YFB2902500)

括高、中、低轨卫星等)等多种组网手段的空天地一体化网络,填补复杂场景下地面网络缺失损毁区域覆盖空洞,增强通信网络保障能力和服务质量^[1]。随着卫星技术、无人机技术的成熟,空天地一体化网络得到了长足发展,在远洋运输、环境监测、抢险救灾等领域发挥了重要作用。与此同时,面对视频回传、环境遥测、远程控制等新兴业务服务质量的差异化需求,如何在空天地一体化网络中实现差异化服务质量保障,成为近年来研究关注的核心问题。

随着微服务技术、网络功能虚拟化技术的发展,网络切片(Network Slicing, NS)技术为解决上述挑战提供了可行的解决途径。网络切片是指在共享的物理网络设施上创建出多个独立的、虚拟化的逻辑网络实例的技术^[2]。每个网络切片是一个独立的端到端网络,具有独立逻辑拓扑和性能特征,以满足特定应用业务的个性化需求。在网络切片中,高效的虚拟网络映射(Virtual Network Embedding, VNE)是保证切片运行的关键。虚拟网络映射是指将逻辑网络映射到共享的物理网络上,以分配相应的物理资源的过程。这是一个在图约束下的大规模背包 NP 难问题^[3]。近年来,虚拟网络映射问题受到了学术界和产业界的广泛关注。例如,参考文献[4]提出了一种卷积神经网络和深度强化学习结合的网络映射方法。参考文献[5]提出了一种基于图神经网络的虚拟网络映射方法。

现有的虚拟网络映射方法通常针对静态网络环境,旨在固定网络拓扑下实现资源的最优分配。然而,空天地一体化网络具有时变特性,其网络资源分布异质且网络拓扑随时间动态变化。将静态网络映射方法直接应用于动态网络时,算法只能针对单个时间点进行最佳资源映射,无法充分考虑动态网络中快照之间的时序关联关系和未来可能发生的网络状态变化。这导致在网络结构发生变化时,需要频繁地重新调整虚拟网络位置,带来大量的资源迁移成本。因此,在设计空天地一体化网络映射方法时,需要考虑不同时间点之间的关联关系,以更好地适应动态网络环境,从而降低迁移成本。

针对这一问题,本文提出了一种空天地一体化网络资源时空关联表征方法,以捕捉动态网络的图结构特征和时间演化特征。该方法利用自注意力机制,在动态网络的图结构邻域和时间序列(时序)邻域两个维度,同时计算节点注意力表征,捕捉节点级细粒度的时空关联关系,实现对空天地一体化网络资源时空关联关系的精准建模^[6]。在此基础上,本文设计了基于时空关联表征的网络资源精准

映射方法,该方法采用节点链路两段虚拟网络映射机制。首先是节点映射过程,采用时空注意力模型对网络时空特征进行建模,并通过全连接神经网络层和 SoftMax 层,对备选映射节点评估并选取最优节点进行映射;其次是链路映射过程,采用弗洛伊德算法计算映射节点间最优路径,完成对链路的映射。在此基础上,采用置信域策略梯度优化(Trust Region Policy Optimization, TRPO),对算法参数进行自适应学习与优化^[7]。研究结果显示,所提出的方法在业务平均收益、请求接收率以及迁移开销等方面的表现优于对照算法。

1 相关工作

参考文献[8]对虚拟网络映射技术做了全面性综述。本节将对虚拟网络映射机制与时变图表征学习方法相关工作分析如下。

1.1 虚拟网络映射机制

虚拟网络映射是一种解决如何合理高效地为虚拟网络请求分配物理网络节点和链路资源的方法。参考文献[9]针对卫星网络的规模、用户数量和服务种类多样等特点,建立了拓扑动态性、服务质量需求和可用资源等约束的虚拟网络映射模型,提出了多约束虚拟网络映射算法。参考文献[10]针对非地面网络中的虚拟网络映射问题,建立了最大化服务收益与最小化功耗与部署成本的映射模型,提出了基于线性松弛的贪心启发式虚拟网络映射算法。参考文献[11]提出了一种针对空天地一体化网络非欧几里得结构的图卷积虚拟网络映射算法。实验结果显示。参考文献[12]针对空天地一体化网络中时延容忍业务和时延敏感业务服务需求,提出了基于时间图的在线虚拟网络映射方法,最大化了服务接收率和底层网络资源利用率。参考文献[13]设计了一种基于深度强化学习和图卷积网络的自适应虚拟网络映射算法,通过并行强化学习框架和多目标奖励函数,该算法在训练中能够快速收敛,解决现有算法在可接收的运行时间内无法提供映射解的问题。参考文献[14]设计了基于强化学习的连续决策虚拟网络映射方法,将相同请求的节点嵌入 Seq2Seq 时间序列模型,以解决传统启发式算法在动态网络结构和环境下难以适用的问题。

1.2 时变图表征学习方法

面向卫星网络、无人机网络、车联网等场景,时变图表征学习方法近年来受到学术界大量关注。参考文献[15]设计了一种动态属性网络嵌入框架 DANE,建立了离线方法来获得一致性的嵌入表示,然后利用矩阵摄动理论以在

线方式保持嵌入结果的新鲜度。参考文献[16]设计了一种名为 **Know-Evolve** 的动态图演化知识网络。该网络学习了实体随时间非线性演化的表示,并将边的发生建模为一个多变量点过程,其中边的强度函数受基于学习的实体嵌入的得分所调节。参考文献[17]设计了一种基于知名图嵌入方法的动态嵌入机制 **Dynnode2vec**。其利用演化随机游走和前一时间戳的嵌入向量初始化当前图嵌入,以解决生成每个时间步的随机游走耗时和不同时间戳的嵌入向量空间不同的问题。参考文献[18]设计了一种动态图表示学习方法 **DynamicTriad**。该方法利用三角组合 (Triad) 作为网络的基本单元,模拟了闭合三角组合从未连接到相互连接的演化过程,从而捕捉了网络形成和演化的基本机制。参考文献[19]针对图数据的微观和宏观动态性特征,提出了动态网络映射方法 **MMDNE**。针对宏观动态性,定义了一个参数化的网络嵌入动力学方程,捕捉动态图内在的演化模式,在更高的结构层次上对网络嵌入进行约束。

2 网络资源动态图模型

本节将对时变物理网络和虚拟网络请求进行建模和描述。

2.1 物理网络时变图模型

在空天地一体化网络中,网络节点和链路状态随时间推移而动态变化,这使得网络模型呈现出典型的时变图特性。为了准确描述时变网络的时间和空间特征,本文将给定时间范围分割为 T 个足够小的离散时间片。不失一般性,本文假设在一个时间片内网络状态保持不变,并采用静态图来描述每个时间片内网络模型,即动态网络的时间快照。如图 1 所示,通过一系列快照的时序累积,时变网络可以呈现为时变图形式,其可以被表示为 $\{G^t(V^t, E^t, C_V^t, C_E^t) | t=1, \dots, T\}$, 其中 $V^t = \{v_1^t, \dots, v_{V^t}^t\}$ 表示 t 时刻节点的集合, $E^t = \{e_1^t, \dots, e_{E^t}^t\}$ 表示 t 时刻链路集合, C_V^t 表示 t 时刻节点 V 的状态属性, C_E^t 表示 t 时刻链路 E 的状态属性。本文将节点 v 的状态属性表征为 $C_V^t = \{\text{Cpu}(v^t), \text{Sub}(v^t), \text{RCpu}(v^t), \text{RBdw}(v^t), \text{Del}(v^t)\}$, 其中 $\text{Cpu}(v^t)$ 代表节点 v^t 的计算能力, $\text{Sub}(v^t)$ 代表节点 v^t 所连接链路中的最大带宽, $\text{RCpu}(v^t)$ 代表 t 时刻节点 v^t 的剩余计算能力, $\text{RBdw}(v^t)$ 代表 t 时刻节点 v^t 所连接链路中最大剩余带宽, $\text{Del}(v^t)$ 代表 t 时刻节点 v^t 所连接链路最大传输时延。此外,本文将链路 e 的状态属性表

征为 $C_e^t = \{\text{Bdw}(e^t)\}$, 其中 $\text{Bdw}(e^t)$ 代表 t 时刻链路 e^t 的可用带宽。

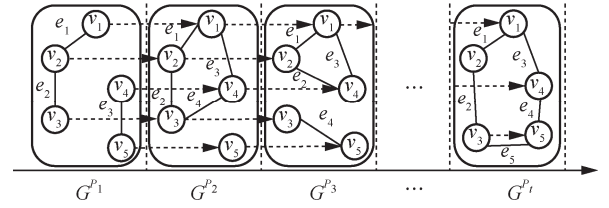


图 1 网络时变图模型

2.2 虚拟网络映射模型

本文将业务 k 的虚拟网络请求建模为 $G^k(V^k, E^k, D_V^k, D_E^k, T^k)$, 其中 $V^k = \{v_1^k, v_2^k, \dots, v_{V^k}^k\}$ 代表第 k 个请求的虚拟节点集合, $E^k = \{e_1^k, e_2^k, \dots, e_{E^k}^k\}$ 第 k 个请求的虚拟链路集合, D_V^k 代表第 k 个请求中虚拟节点 V 的资源需求, D_E^k 代表第 k 个虚拟请求中链路 E 的资源需求, T^k 代表第 k 个虚拟请求的持续时间。本文设置 $D_V^k = \{\text{Cpu}(v^k)\}$ 和 $D_E^k = \{\text{Bdw}(e^k)\}$ 。因此,虚拟网络映射问题可以被形式化描述为: $M: G^k(V^k, E^k, D_V^k, D_E^k, T^k) \rightarrow G^t(V^t, E^t, C_V^t, C_E^t) | t=t, \dots, t+T^k$ 。本文定义了节点映射度量 $\alpha_{v^k}^{v^t}$ 和链路映射度量 $\kappa_{e^k}^{e^t}$ 来表示节点与链路是否映射成功,其可被表示为

$$\alpha_{v^k}^{v^t} = \begin{cases} 1, & v^k \rightarrow v^t \\ 0, & \text{其他} \end{cases} \quad (1)$$

$$\kappa_{e^k}^{e^t} = \begin{cases} 1, & e^k \rightarrow e^t \\ 0, & \text{其他} \end{cases} \quad (2)$$

此外,虚拟网络映射过程中,节点资源映射需要满足如下约束

$$\sum_{k \in K} (\alpha_{v^k}^{v^t} \times D_V^k) \leq C_V^t \quad (3)$$

链路资源映射需要满足如下约束

$$\sum_{k \in K} (\kappa_{e^k}^{e^t} \times D_E^k) \leq C_E^t \quad (4)$$

3 基于时空关联表征的网络资源精准映射方法

本节设计了基于时空关联表征的空天地一体化网络资源精准映射方法。如图 2 所示,该算法包含 4 个部分,分别是空天地一体化网络的空间注意力模块、时间注意力模块、映射模块、置信域策略梯度优化模块。其中,空间注意力模块用于提取空天地一体化网络资源图特征的关联关系,时间注意力模块用于提取空天地一体化网络资源时间维度下的关联关系,映射模块用于选择

$$\delta_{e^p}^t = \begin{cases} 1, & \sum_{e^p \in E^{p_t}} |f_{e^{p+1}}^{e^p} - f_{e^p}^{e^p}| > 0, \forall t \in T^{R_k} \\ 0, & \text{其他} \end{cases} \quad (9)$$

3.2 空间注意力模块

本文采用图注意力网络 (Graph Attention Network, GAT) 提取网络状态的空间结构关系^[21]。GAT 是一种用于处理图结构关系数据的模型, 通过注意力机制实现节点特征的动态更新, 以捕捉节点之间复杂关系。其核心思想是通过可学习的注意力权重, 使得每个节点能够根据其与其邻居节点的相关性动态地调整其重要性, 通过学习得到的权重矩阵实现信息的传递与聚合, 从而实现了图结构的有效建模^[22]。图注意力网络由多个相互连接的图注意力层组成。每个图注意力层的初始输入是网络快照 G^p 中节点属性编码向量 $\{h_v^0, \forall v \in V\}$ 集合, 通常采用 one-hot 向量编码。该层将输入转换为一个新的节点表示集合 $\{h_v^1, \forall v \in V\}$, 并将这些输出向量作为输入传递给下一层的图注意力层。图注意力层的更新公式可被表示为

$$h_v^{(l+1)} = \sigma \left(\sum_{u \in V} \alpha_{vu}^{(l)} \cdot W^{(l)} \cdot h_v^{(l)} \right) \quad (10)$$

其中, $h_v^{(l+1)}$ 代表第 v 个节点在第 $l+1$ 层的聚合特征表示, σ 代表激活函数 (通常使用 ReLU 函数), $\alpha_{vu}^{(l)}$ 代表节点 v 对节点 u 的注意力权重, $W^{(l)}$ 代表第 l 层的权重矩阵, $h_v^{(l)}$ 代表第 v 个节点在第 l 层的聚合特征表示。注意力权重 $\alpha_{vu}^{(l)}$ 是通过计算两个节点之间的注意力分数来确定的, 可被表示为

$$\alpha_{vu}^{(l)} = \text{SoftMax}(e_{vu}^{(l)}) = \frac{\exp(e_{vu}^{(l)})}{\sum_{u=1}^V \exp(e_{vu}^{(l)})} \quad (11)$$

其中, $e_{vu}^{(l)}$ 代表节点 v 和节点 u 之间的注意力分数, 可被表示为

$$e_{vu}^{(l)} = \text{LeakyReLU}(a^{(l)T} [W^{(l)} h_v \| W^{(l)} h_u]) \quad (12)$$

其中, LeakyReLU 代表激活函数, $[W^{(l)} h_v \| W^{(l)} h_u]$ 代表 $W^{(l)} h_v$ 与 $W^{(l)} h_u$ 的串联向量, $a^{(l)}$ 表示邻居节点与目标节点间相关度参数。注意, $a^{(l)}$ 是一个可学习的权重向量, 通过在学习过程中不断调整来确定每个节点的注意力分布。因此, 注意力权重表达式可被重新表述为

$$\begin{aligned} \alpha_{vu}^{(l)} &= \text{SoftMax}(e_{vu}^{(l)}) \\ &= \frac{\exp(L(a^{(l)T} [W h_v \| W x_u]))}{\sum_{u \in V} \exp(L(a^{(l)T} [W x_v \| W x_u]))} \end{aligned} \quad (13)$$

通过上述过程, 每个节点都会根据其与其相邻节点的关系以及关系的重要程度来更新自己的特征表示, 实现了信

息的交互, 并通过循环迭代实现空间信息传递, 实现对复杂空间关联结构的精准建模。本文定义空间注意力模块的输出向量为 $\{h_v^{\text{out}}, \forall v \in V\}$ 。

3.3 时间注意力模块

为了刻画动态网络中网络状态的时间关联关系, 本文设计了时间注意力模块^[23]。时间注意力方法是一种用于捕捉动态网络中节点变化模式的神经网络结构。该方法的输入是动态网络的时间快照中节点 v 在不同时间片的表征序列 $\{h_v^{(\text{out}), t_1}, h_v^{(\text{out}), t_2}, \dots, h_v^{(\text{out}), t_N}\}$, 其中 N 是时间序列长度; 并通过时间注意力机制, 学习输入数据不同时间步之间的关联关系, 捕捉节点在时间变化过程中的重要特征, 从而生成新的表示序列 $\{z_v^1, z_v^2, \dots, z_v^N\}$ 。时序注意力层的更新公式可被表示为:

$$z_v = \beta_v (H_v^{\text{out}} W_v) \quad (14)$$

其中, $H_v^{\text{out}} W_v$ 代表 H_v^{out} 在数值 (Value) 矩阵 W_v 上的投影, β_v 代表注意力权重矩阵, 通常采用 SoftMax 函数, 其计算公式为

$$\beta_v = \frac{\exp(e_v^{t_k})}{\sum_{z=1}^T \exp(e_v^{t_z})} \quad (15)$$

其中, $e_v^{t_k}$ 代表节点 v 在 t 时刻和 k 时刻之间的注意力分数, 用于衡量节点在不同时间步之间的关联关系, 其计算更新公式可被表示为

$$e_v^{t_k} = \left(\frac{((H_v^{\text{out}} W_q)(H_v^{\text{out}} W_k)^T)_{tk}}{\sqrt{F}} + M_{tk} \right) \quad (16)$$

其中, $H_v^{\text{out}} W_q$ 代表 H_v^{out} 在查询 (Query) 矩阵 W_q 上的投影, $H_v^{\text{out}} W_k$ 代表 H_v^{out} 在键 (Key) 矩阵 W_k 上的投影, M_{tk} 代表可学习的偏置项, 防止模型注意到未来的时间步, 其更新公式可被表示为

$$M_{tk} = \begin{cases} 0, & i < j \\ -\infty, & \text{其他} \end{cases} \quad (17)$$

在自注意力机制中, 本文通过计算查询矩阵 $H_v^{\text{out}} W_q$ 与键矩阵 $H_v^{\text{out}} W_k$ 投影将输入序列中的元素映射到一个抽象的空间, 然后对数值矩阵 $X_v W_v$ 中的值进行加权求和从而得到最终的自注意力输出, 实现对时序关联关系的建模。此外, 本文引入位置编码来帮助模型理解输入序列中元素的位置顺序, 其计算公式可被表示为

$$\text{PE}_{(\text{pos}, 2i)} = \sin \left(\frac{\text{pos}}{10\,000^{\frac{2i}{d_{\text{model}}}}} \right) \quad (18)$$

$$PE_{(\text{pos}, 2i+1)} = \cos\left(\frac{\text{pos}}{10\,000 \frac{2i}{d_{\text{model}}}}\right) \quad (19)$$

其中, $PE_{(\text{pos}, 2i)}$ 和 $PE_{(\text{pos}, 2i+1)}$ 分别表示位置 pos 上奇偶维度的位置编码向量, d_{model} 表示模型的维度。基于上述位置信息建模, 本文将输入表征矩阵重新表示为 $\{h_v^{(\text{out}), t_1} + PE_1, h_v^{(\text{out}), t_2} + PE_2, \dots, h_v^{(\text{out}), t_N} + PE_N\}$ 。

3.4 置信域策略梯度优化模块

本小节将运用强化学习来优化空间注意力模块、时间注意力模块和映射模块的参数。注意, 本文中映射模块主要包含一个全连接神经网络层和一个 SoftMax 层, 用于对备选映射节点进行打分 (本文将不在此处详细描述)。为了方便后续描述, 本文将空间注意力模块和时间注意力模块的参数及映射模块参数合并为一个参数集合, 记为 W 。其中, 策略函数可表示为 $\pi(A|S; W)$, 优化目标可表示为

$$\max_W J(W) = \max_W E_S[V_\pi(S; W)] \quad (20)$$

其中, $V_\pi(S; W)$ 代表策略 π 下的值函数, 可表示为 $V_\pi(S; W) = \sum_a \pi(a|S; W) \cdot Q_\pi(s, a)$ 。为了优化该目标, 传统随机梯度上升方法通过让参数值不断逼近梯度最大值来最大化预期累积奖励。然而, 传统的随机梯度上升方法难以确保单调收敛。这是受到梯度估计的噪声和参数更新的不确定性影响, 从而导致训练过程的不稳定。特别是在虚拟网络映射问题中, 奖励信号的传播步长较长, 需要等到虚拟业务完全服务完毕后才能对单个动作的奖励进行评估。这导致了目标奖励信号与动作之间的关联噪声累积大, 传统随机梯度上升方法难以有效收敛。针对这一问题, 本文采用了置信域策略梯度优化算法。该算法通过引入置信域约束来限制参数更新的幅度, 确保每次更新都处于可控范围内, 从而避免了传统策略梯度优化算法中出现的剧烈波动和不稳定性。

置信域策略梯度优化算法主要过程是通过构造 W_{old} 在置信域内的近似函数 $L(W|W_{\text{old}})$, 并通过计算 $L(W|W_{\text{old}})$ 梯度来近似 $J(W)$ 相对参数的梯度。其参数更新过程可被表示为

$$W_{\text{new}} \leftarrow \arg \max_W L(W|W_{\text{old}}); \quad \text{s.t. } W \in \mathcal{N}(W_{\text{old}}) \quad (21)$$

其中, $\mathcal{N}(W_{\text{old}})$ 为 W_{old} 的置信域, 其是 W_{old} 的一个邻域约束, 用于限制参数更新的幅度, 以确保每次更新都在一个可控的范围内。通常可以使用 KL 散度来衡量 $\mathcal{N}(W_{\text{old}})$ 置信域标准, 其数学表达式为

$$\frac{1}{n} \sum_{i=1}^n \text{KL}[\pi(\cdot|s_i; W_{\text{old}}) \| \pi(\cdot|s_i; W)] < \Delta \quad (22)$$

即通过限制策略 W_{old} 与 W 之间的 KL 散度, 以确保参数更新的大小在一个可控的范围内。如式(22)所示, 置信域策略梯度优化算法的核心是构造置信域内 $J(W)$ 的近似函数 $L(W|W_{\text{old}})$ 。本文首先对优化函数进行改写

$$\begin{aligned} V_\pi(S) &= \sum_a \pi(a|S; W) \cdot Q_\pi(s, a) \\ &= \sum_a \pi(a|S; W_{\text{old}}) \cdot \frac{\pi(a|S; W)}{\pi(a|S; W_{\text{old}})} \cdot Q_\pi(s, a) \\ &= E_{A \sim \pi(\cdot|S; W_{\text{old}})} \left[\frac{\pi(a|S; W)}{\pi(a|S; W_{\text{old}})} \cdot Q_\pi(s, a) \right] \end{aligned} \quad (23)$$

因此, 优化目标函数可重构为

$$\max_W J(W) = E_{S, A} \left[\frac{\pi(A|S; W)}{\pi(A|S; W_{\text{old}})} \cdot Q_\pi(S, A) \right] \quad (24)$$

为了计算该期望, 本文对期望进行蒙特卡洛近似

$$L(W|W_{\text{old}}) = \frac{1}{n} \sum_{i=1}^n \frac{\pi(a_i|s_i; W)}{\pi(a_i|s_i; W_{\text{old}})} \cdot Q_\pi(s_i, a_i) \quad (25)$$

其中, $\{s_1, a_1, r_1, \dots, s_n, a_n, r_n\}$ 为策略的动作轨迹。此外, 本文将 $Q_\pi(s_i, a_i)$ 近似为折现期望收益

$$Q_\pi(s_i, a_i) = r_i + \gamma r_{i+1} + \gamma^2 r_{i+2} + \dots + \gamma^n r_{i+n} \quad (26)$$

算法 1: 基于时空关联表征的网络资源精准映射方法

1: 初始化物理网络环境, 初始化虚拟网络请求; 初始化神经网络参数;

2: 遍历 epoch = 1 到 max - Num ;

3: 数量 Num 设为 0;

4: 遍历虚拟网络请求集合中的请求 G^{R_k} do:

5: 遍历请求 G^{R_k} 中的节点 V^{R_k} do:

6: 使用时空注意力算法与映射模型寻找可嵌入节点;

7: 结束

8: 如果所有节点 V^{R_k} 嵌入成功 then:

9: 遍历在请求 G^{R_k} 中的链路 E^{R_k} do:

10: 使用弗洛伊德算法寻找可嵌入链路;

11: 结束

12: 结束

13: 如果节点 V^{R_k} 和链路 E^{R_k} 在 T^{R_k} 时间都能成功嵌入: 则计算置信梯度;

14: 否则: 删除置信梯度;

15: 如果 Num 达到 batchSize, 则:

16: 参数更新;

17: 结束

18: 结束

在此基础上,本文设计了基于时空关联表征的网络资源精准映射方法(算法 1)。需要注意的是,在运行过程中,如果节点或链路在时间 T^R 内的任一时刻映射失败,则认为映射失败,不参与奖励值反馈。

4 实验结果与分析

本节将对所提方法进行实验仿真,并对结果进行分析。

4.1 仿真设置

本文模拟了一个包含 100 个节点和 500 条链路的网络。其中,每个节点的计算资源 $Cpu(v)$ 容量服从于 (50,100) 之间的均匀分布,带宽 $Bdw(v)$ 服从 (20,50) 之间的均匀分布。此外,本文模拟了 1 000 组虚拟网络请求集合,其中,每个虚拟网络中节点的数量服从 (2,10) 之间的均匀分布,计算资源服从 (0,50) 之间的均匀分布,每个虚拟链路的带宽资源遵循 (0,50) 之间的均匀分布,链路之间的连接概率设定为 0.5。虚拟请求的到达时间遵循泊松分布,其中,每 100 个时间单位内会到达 4 个请求。每个请求的持续时间呈指数分布,平均持续时间为 1 000 个时间单位。此外,在本实验中,设置了 3 种对比算法,分别是基于启发式的 NodeRank 算法^[24]、基于强化学习的 PG-CNN 算法^[25]和基于图神经网络的 GCN-VNE 算法^[26]。

4.2 收敛性分析

在本实验中对所提方法的收敛性进行分析。如图 3 所示,随着训练步数的增加,本文所提方法快速收敛。这是因为本文所提方法采用置信域策略梯度优化机制。该机制通过引入置信域约束来限制参数更新的幅度,即引入一个信任区域,确保每次更新都处于可控范围,从而避免了传统策略梯度优化算法中出现的剧烈波动和不稳定性。

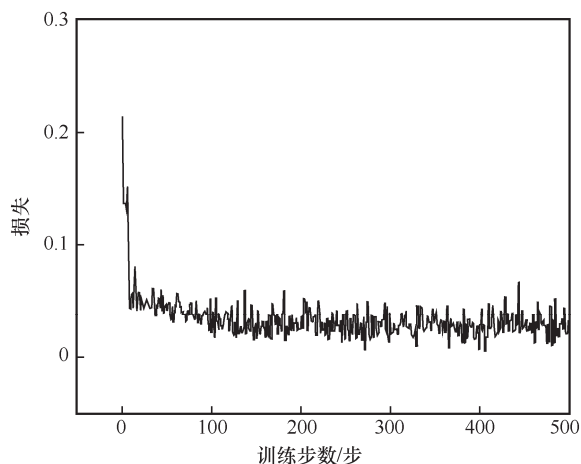


图 3 收敛性分析

4.3 长期平均接收率与平均长期收益分析

在本实验中,对比 4 种映射算法的长期平均接收率 A_{cp} (如图 4 所示)与平均长期收益 (如图 5 所示)。长期平均接收率的计算方式为 $A_{cp} = \sum_{t=0}^T A_{cp} / \sum_{t=0}^T All$, 其中, A_{cp} 表示成功映射的虚拟网络个数, All 表示请求总数。图 4、图 5 中, NodeRank 算法表现出最差的性能,这是因为启发式算法难以有效捕捉网络动态性特征。基于强化学习的 PG-CNN 算法表现出较好的性能,然而由于缺少对空天地一体化网络时空动态性特征的建模,导致性能损失。GCN-VNE 能够有效学习到空间特征,但是缺少时序特征建模。本文所提方法在网络的图结构邻域和时间序列邻域两个维度同时计算动态节点注意力表征,能够捕捉节点级细粒度的图空间结构和时序结构关联关系,从而提高虚拟任务的平均接收率。

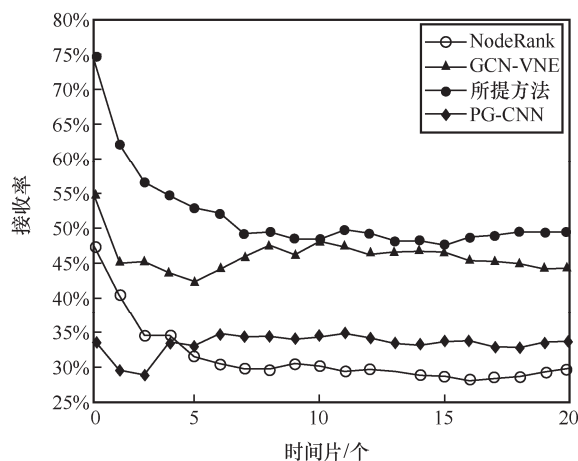


图 4 长期平均接收率对比

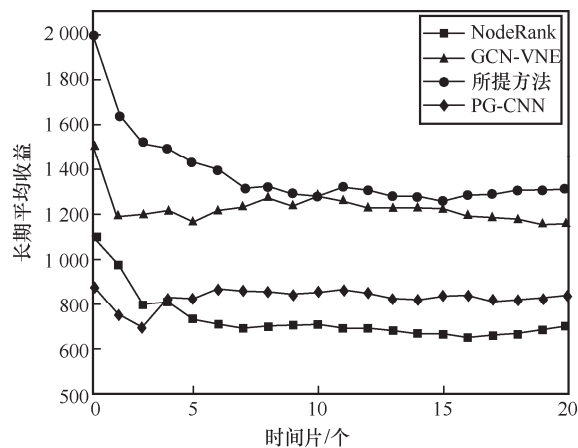


图 5 长期平均长期收益对比

4.4 迁移成本与收益成本比分析

在本实验中，对比4种映射算法的平均迁移成本（如图6所示）与收益成本比（如图7所示），可以看出，传统方法在静态时间片上进行最佳资源映射，而未能考虑快照之间的时序关联关系。这会导致在网络结构变化时，需要频繁地重新调整虚拟网络的位置，从而产生大量的迁移开销。相比之下，本文所提方法能够有效捕捉空间结构和时序结构关联关系，从而有效降低迁移成本。在此基础上，在同等的性能收益条件下，收益成本比有效提高。

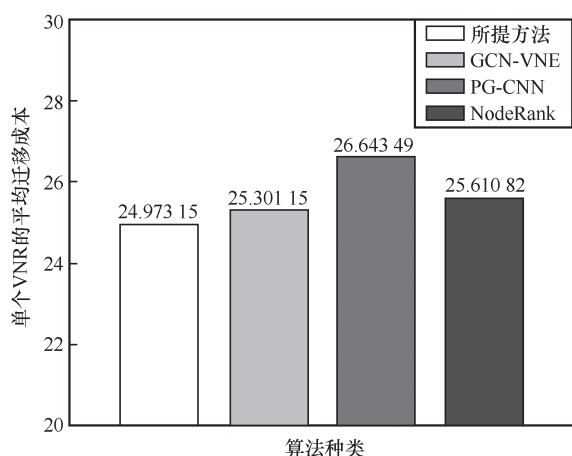


图6 平均迁移成本对比

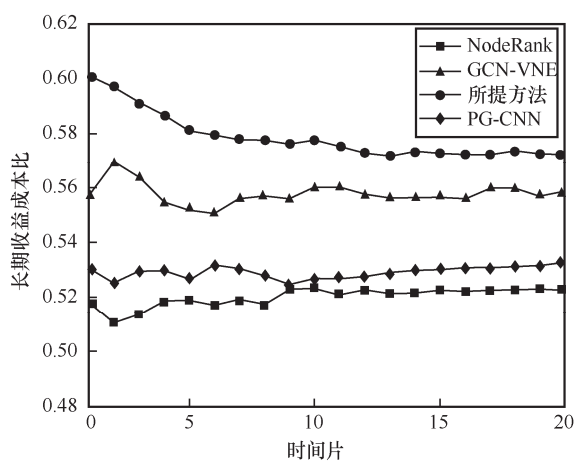


图7 收益成本比

参考文献：

- [1] 买天乐. 复杂场景下异构融合网络智能组网关键技术[D]. 北京: 北京邮电大学, 2023.
MAI T L. Key technologies for intelligent networking of heterogeneous integrated networks in complex environments[D]. Beijing: Beijing University of Posts and Telecommunications, 2023.
- [2] 项弘禹, 肖扬文, 张贤, 等. 5G 边缘计算和网络切片技术[J]. 电信科学, 2017, 33(6): 54-63.
- [3] XIANG H Y, XIAO Y W, ZHANG X, et al. Edge computing and network slicing technology in 5G[J]. Telecommunications Science, 2017, 33(6): 54-63.
- [4] WU H T, ZHOU F, CHEN Y J, et al. On virtual network embedding: paths and cycles[J]. IEEE Transactions on Network and Service Management, 2020, 17(3): 1487-1500.
- [5] DOLATI M, HASSANPOUR S B, GHADERI M, et al. DeepViNE: virtual network embedding with deep reinforcement learning[C]//Proceedings of the IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2019: 879-885.
- [6] RKHAMI A, QUANG P T A, HADJADJ-AOUL Y, et al. On the use of graph neural networks for virtual network embedding[C]//Proceedings of the 2020 International Symposium on Networks, Computers and Communications (ISNCC). Piscataway: IEEE Press, 2020: 1-6.
- [7] SANKAR A, WU Y H, GOU L, et al. Dynamic graph representation learning via self-attention networks[EB]. 2018.
- [8] SCHULMAN J, LEVINE S, MORITZ P, et al. Trust region policy optimization[EB]. 2015.
- [9] FISCHER A, BOTERO J F, BECK M T, et al. Virtual network embedding: a survey[J]. IEEE Communications Surveys & Tutorials, 2013, 15(4): 1888-1906.
- [10] YANG D, LIU J, ZHANG R, et al. Multi-constraint virtual network embedding algorithm for satellite networks[C]//Proceedings of the GLOBECOM 2020 - 2020 IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-6.
- [11] MAITY I, VU T X, CHATZINOTAS S, et al. D-ViNE: dynamic virtual network embedding in non-terrestrial networks[C]//Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE Press, 2022: 166-171.
- [12] CHEN N, SHEN S G, DUAN Y X, et al. Non-euclidean graph-convolution virtual network embedding for space-air-ground integrated networks[J]. Drones, 2023, 7(3): 165.
- [13] LIU B, ZHANG T, ZHANG L M, et al. Online virtual network embedding for both the delay sensitive and tolerant services in SDN-enabled satellite-terrestrial networks[C]//Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE Press, 2023: 1-6.
- [14] YAN Z X, GE J G, WU Y L, et al. Automatic virtual network embedding: a deep reinforcement learning approach with graph

- convolutional networks[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(6): 1040-1057.
- [14] YAO H P, MA S H, WANG J J, et al. A continuous-decision virtual network embedding scheme relying on reinforcement learning[J]. IEEE Transactions on Network and Service Management, 2020, 17(2): 864-875.
- [15] LI J D, DANI H, HU X, et al. Attributed network embedding for learning in a dynamic environment[C]//Proceedings of the Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. New York: ACM, 2017: 387-396.
- [16] TRIVEDI R, DAI H J, WANG Y C, et al. Know-evolve: deep temporal reasoning for dynamic knowledge graphs[C]//Proceedings of the Proceedings of the 34th International Conference on Machine Learning - Volume 70. New York: ACM, 2017: 3462-3471.
- [17] MAHDAVI S, KHOSHRAFTAR S, AN A J. dynnode2vec: scalable dynamic network embedding[C]//Proceedings of the 2018 IEEE International Conference on Big Data (Big Data). Piscataway: IEEE Press, 2018: 3762-3765.
- [18] ZHOU L K, YANG Y, REN X, et al. Dynamic network embedding by modeling triadic closure process[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 1-5.
- [19] LU Y F, WANG X, SHI C, et al. Temporal network embedding with micro- and macro-dynamics[C]//Proceedings of the Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: ACM, 2019: 469-478.
- [20] LIM H K, ULLAH I, HAN Y H, et al. Reinforcement learning-based virtual network embedding: a comprehensive survey[J]. ICT Express, 2023, 9(5): 983-994.
- [21] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB]. 2017.
- [22] BRODY S, ALON U, YAHAV E. How attentive are graph attention networks? [EB]. 2021.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [24] CHENG X, SU S, ZHANG Z B, et al. Virtual network embedding through topology-aware node ranking[J]. ACM SIGCOMM Computer Communication Review, 2011, 41(2): 38-47.
- [25] YAO H P, LIU H W, ZHANG P Y, et al. A learning-based approach to intra-domain QoS routing[J]. IEEE Transactions on Vehicular Technology, 2020, 69(6): 6718-6730.
- [26] MA S H, YAO H P, MAI T L, et al. Graph convolutional network aided virtual network embedding for Internet of thing[J]. IEEE Transactions on Network Science and Engineering, 2023, 10(1): 265-274.

[作者简介]



买天乐 (1994—), 男, 北京理工大学博士后, 在 IEEE TCC、IEEE TON、IEEE TSC 等学术期刊发表多篇学术论文, 相关研究成果获 IEEE ICC、IEEE IWCMC 最佳论文奖等, 主要研究方向为空天地一体化网络、无人集群网络、网络人工智能、多智能体系统等。



姚海鹏 (1983—), 男, 北京邮电大学教授, 在学术期刊和会议上发表论文 150 余篇, 担任 IEEE TMC 和 IEEE T-SUSC 期刊 Associate Editor, 主要研究方向为未来网络架构、网络人工智能、空天地一体化网络、网络资源分配和专用网络等。



忻向军 (1969—), 男, 北京理工大学教授, 在学术期刊和会议上发表论文 100 余篇, 主要研究方向为光通信、卫星通信等。



杨景凯 (1996—), 男, 北京邮电大学博士生, 主要研究方向为未来网络架构、网络资源分配和多智能体系统。



靳辰朗 (2000—), 女, 北京邮电大学博士生, 主要研究方向为未来网络架构、资源优化和博弈论等。