

基于注意力神经网络的对地观测卫星星上自主任务规划方法

彭 双 伍江江 陈 浩 杜 春 李 军

国防科技大学电子科学学院 长沙 410073

(pengshuang08@nudt.edu.cn)

摘 要 星上自主任务规划是对地观测卫星自主运行的关键技术之一,近年来得到了研究人员的高度关注。考虑到星上计算资源有限,以及星上任务、资源动态变化等特点与挑战,现有研究主要采用启发式搜索算法对卫星星上自主任务规划问题进行求解,但这类算法还有待进一步优化。文中首先构建了一种新的观测任务序贯决策框架。基于该框架,对地观测卫星可以实时决策要执行的观测任务,而无须预先生成任何观测方案。然后,将注意力机制和循环神经网络相结合,设计了观测任务决策模型、任务特征表示方法以及模型训练方法,提出了一种基于注意力神经网络的观测任务序贯算法;最后,基于多组随机数据对所提算法、两种深度学习算法以及两种启发式在线搜索算法进行了比较。实验结果表明,所提方法的平均响应时间不到已有深度学习算法的 1/5,收益误差远低于启发式搜索算法,证实了所提方法的可行性和有效性。

关键词: 对地观测卫星;星上自主任务规划;序贯决策;循环神经网络;注意力机制

中图法分类号 TP391

Satellite Onboard Observation Task Planning Based on Attention Neural Network

PENG Shuang, WU Jiang-jiang, CHEN Hao, DU Chun and LI Jun

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

Abstract Satellite onboard autonomous task planning is one of the key technologies for the operation of earth observation satellites, which has received great attention from researchers in recent years. Considering the limited computing resources, as well as the dynamic changes of observation tasks and resource onboard, the heuristic search algorithms are mainly used to solve the satellite onboard task planning problem, and the optimization of solution needs to be improved. Firstly, a new sequential decision-making framework for observation tasks is constructed in this paper. Based on this framework, an earth observation satellite can decide the observation task to be performed in real-time, without generating a plan in advance. Then, an observation task decision model based on attention mechanism, and the corresponding input feature representation method and model training method are designed. An observation task sequence algorithm based on attention neural network is proposed. Finally, based on a set of random data, the performance of the proposed algorithm, two deep learning algorithms and two heuristic online search algorithms are compared. Experimental results show that the response time of the proposed method is less than one-fifth of the existing deep learning algorithm, and the profit gap is much smaller than that of the heuristic search algorithms, which confirm the feasibility and effectiveness of our method.

Keywords Earth observation satellite, Satellite onboard autonomous task planning, Sequential decision-making, Recurrent neural network, Attention mechanism

1 引言

对地观测卫星通常由地面规划系统集中管理和控制^[1-2]。然而,由于星地通信时间窗口的限制以及通信时延的影响,地面运控系统往往难以快速对星上能量和存储等资源的使用与预期不符、载荷设备暂时失效以及自主探测到高价值目标等紧急突发情况进行动态响应^[3]。而卫星星上自主任务规划能够有效弥补上述不足,减少对地面系统的依赖,增强对卫星

运行过程中星上任务、资源等动态变化的快速响应能力,得到了各航天大国和研究人员的高度关注,是未来卫星任务规划技术发展的一个重要方向^[4]。

目前,卫星星上自主任务规划问题的研究已经取得了一定的成果。总体上,研究人员普遍将该问题建模为优化问题,基于滚动/连续规划框架动态迭代求解^[5-7]。在该框架下,规划器根据星上实时获取的任务、资源及设备状态信息,动态地对未来短期范围内的观测任务进行规划,并根据其变化对

到稿日期:2021-05-13 返修日期:2021-09-08

基金项目:国家自然科学基金(62106276);湖南省自然科学基金(2020JJ4103)

This work was supported by the National Natural Science Foundation of China(62106276) and Natural Science Foundation of Hunan Province (2020JJ4103).

通信作者:陈浩(hchen@nudt.edu.cn)

观测方案进行迭代更新。该框架已在 EO-1 对地观测卫星、Rosetta 深空探测器等多个航天器的规划调度系统中得到应用^[5-7]。除上述自主规划系统外,还有对星上自主任务规划算法的研究,如 Liu 等^[8]将多种任务排序规则与轮盘赌思想相结合,提出了一种迭代贪婪搜索算法;Chu 等^[9]将多种优化策略与分支定界算法相结合,提出了一种随时分支定界算法来调度目标识别任务;Li 等^[10]设计了两种规划时机驱动策略和两种启发式算法来调度常规任务和紧急任务。可以看出,上述研究主要采用迭代修复方法对规划方案进行动态生成和更新,规划方案的优化效果很大程度上取决于修复策略。但是,设计出一种好的修复策略无疑是非常困难的,因此往往难以得到优化的规划方案。

随着机器学习技术的迅速发展,部分研究人员尝试将机器学习算法引入到规划问题中,利用从历史规划结果和试错中学习到的知识来指导规划,以进一步提高算法的求解效率及优化规划结果^[11],如 Wang 等^[12]将卫星任务规划问题映射为动态随机背包问题,采用强化学习算法来判断动态到达的每个观测任务是否应插入到观测方案中。Li 等^[13]和 Lu 等^[14]利用神经网络计算每个任务的调度优先级,并与启发式搜索算法相结合来对问题进行求解;Zhao 等^[15]则提出了基于强化学习的两阶段神经组合优化方法来解决敏捷成像卫星任务调度问题。上述研究实验结果表明,结合机器学习模型的任务规划方法在优化性能上通常要好于基于启发式规则的搜索算法,但上述方法均要求提前制定好初始方案,且对任务、资源变化的响应时间会随着动态变化任务或受影响任务数量的增加而增加。

文献[16]提出了基于循环神经网络的观测任务序贯决策算法。基于该算法,对地观测卫星可以实时决策要执行的任务而不受任务数量规模及其任务、资源变化程度的影响,不需要提前制定方案,有效提高了卫星应对突发情况的快速响应能力,但其求解效率相对较低,具有较大的提升空间。本文在文献[16]工作的基础上,对观测任务序贯决策框架和观测任务决策模型进行了重新设计,进一步提升了任务决策效率。其主要贡献如下:

(1)设计了新的观测任务序贯决策框架,将计算耗时的任务特征提取部分与任务调度结果决策部分分离,对同一轨道圈次内的所有任务特征进行批量提取和重复使用,有效减少了任务特征的计算次数,提升了决策效率;

(2)针对新框架,设计了注意力机制与循环神经网络相结合的观测任务决策模型、观测任务输入特征表示方法以及模型训练方法,提出了基于注意力神经网络的观测任务序贯决策算法;

(3)利用多组随机测试用例验证了所提方法的可行性与高效性。

本文第2节对卫星上自主任务规划问题进行了描述,明确了约束条件和优化目标;第3节建立了对应的数学模型;第4节设计了观测任务序贯决策框架,明确了任务决策的具体流程;第5节提出了基于注意力神经网络的观测任务决策模型、任务特征表示方法以及模型的训练方法;第6节给出了实验结果并对其进行了解析;最后总结全文。

2 问题描述

卫星上自主任务规划问题是一类典型的组合优化问题,其目标是在不违反任何约束条件的前提下从候选的任务集合中挑选出一部分任务来执行,以最大化整体效益。本文将具有星上自主任务规划能力的对地观测卫星称为自主运行卫星(Autonomous Earth Observation Satellite, AEOS)。

如图1所示,AEOS绕地球飞行,当地面目标进入观测范围时,其需要自主判断是否对地面目标进行成像(即对地观测);当AEOS进入地面站的数据接收范围内时,其需要自主决策是否将星载存储器中的数据下传至地面(即对地数传)。卫星在对地观测或对地数传过程中均会消耗星载电池中储存的能量。卫星对地观测后需要将生成的遥感影像保存到星载存储器中,会占用一定的存储空间;而卫星对地数传后会删除星载存储器中已下传的遥感影像数据,释放相应的存储空间。

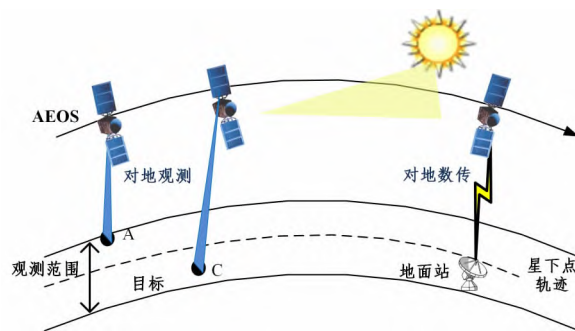


图1 卫星上自主任务规划示意图^[16]

Fig. 1 Illustration of satellite onboard autonomous task planning^[16]

由于同一地面站无法在同一时刻同时接收两颗卫星下传的数据,为了避免多颗卫星因抢占同一地面站资源造成的数据接收冲突,卫星对地数传活动已由地面规划系统提前确定且不会改变,因此本文仅对卫星对地观测活动进行规划,即仅解决卫星上观测任务规划问题(Satellite Onboard Observation Task Planning, SOOTP)。

2.1 约束条件

(1)任务切换约束

AEOS对地面目标进行成像时对观测角度有着明确的要求,当卫星由当前观测任务切换到下一个观测任务时,需要花费一定的时间来对卫星姿态进行调整,以满足下一个任务的观测要求。因此,任意两个连续执行的观测任务之间的时间间隔不能小于最短切换时间,其可根据相邻两个观测任务的观测角度和卫星姿态调整运动方程计算得到。

(2)能量约束

AEOS在执行任务过程中会消耗星载电池中储存的能量,当其处于光照区且未执行任务时将太阳能帆板指向太阳来补充能量。受星载电池容量的限制,其能量状态在任何时候都不能高于电池容量上限;为了保证星载电池的使用寿命,其能量状态不能低于设定的下限。

(3)存储约束

AEOS执行观测任务时生成的遥感影像保存在星载存储器中,占用了存储空间;对地数传后,星载存储器中的遥感影像数据被下传至地面,占用的存储空间得到释放。与能量约束类似,星载存储器中的数据状态在任何时候都不能

超过存储器容量的限制。

2.2 优化目标

本文以最大化观测收益为目标,即在不违反任务切换、能量、存储等约束条件的前提下,使规划安排观测任务的累计收益值最大化。

3 数学模型

根据问题描述,我们给出了要使用的公式符号,如表 1 所列,并建立相应的数学模型。

表 1 符号定义

Table 1 Definition of symbols

符号	含义
T_{sta}	任务规划开始时间
T_{end}	任务规划结束时间
O	规划时段 $[T_{sta}, T_{end}]$ 内,卫星绕地球飞行轨道圈数
M_k	规划时段 $[T_{sta}, T_{end}]$ 的第 k 个轨道圈次内观测任务集合, $1 \leq k \leq O$
M	规划时段 $[T_{sta}, T_{end}]$ 内所有观测任务集合, $M = \{M_1, \dots, M_O\}$
m_i	观测任务, $m_i \in M$
ta_i	观测任务 m_i 的成像开始时间
te_i	观测任务 m_i 的成像结束时间
an_i	观测任务 m_i 成像时的侧摆角度
b_i	观测任务 m_i 的收益值,取值范围 $[1-40]$
o_i	观测任务 m_i 的轨道圈号,若 $o_i = k$,则有 $m_i \in M_k$
e_i	执行观测任务 m_i 消耗的能量, $e_i > 0$
d_i	执行观测任务 m_i 占用的存储空间, $d_i > 0$
e_{ij}	任务 m_i 和任务 m_j 之间的能量变化量,包括卫星帆板充电时的能量补充量和姿态调整时消耗的能量
T_{sij}	任务 m_i 和任务 m_j 之间的最短切换时间,取决于它们的观测角度
$e(t)$	t 时刻星载电池的能量状态,即剩余的能量
$d(t)$	t 时刻星载存储器的存储状态,即剩余的存储空间
E_0	T_{sta} 时刻,星载电池剩余的能量
D_0	T_{sta} 时刻,星载存储器剩余的存储空间
E_{min}	星载电池的最低电量
E_{max}	星载电池的最大可存储电量
D_{max}	星载存储器的最大存储容量
x_i	决策变量, $x_i \in \{0, 1\}$, $x_i = 1$ 表明任务 m_i 被安排,否则被舍弃
x_{ij}	变量, $x_{ij} \in \{0, 1\}$,若 m_i 和 m_j 是两个连续被安排执行的任务,则 $x_{ij} = 1$,否则 $x_{ij} = 0$,即
	$x_{ij} = \begin{cases} 1, & \sum_{p=i}^j x_p = x_i + x_j = 2 \\ 0, & \text{其他} \end{cases}$

根据问题描述及给定的符号,卫星星上自主任务规划问题建模如下。

优化目标为:

$$\max \{ \sum_{m_i \in M} x_i * b_i \} \quad (1)$$

约束条件为:

$$e(te_0) = E_0 \quad (2)$$

$$d(te_0) = D_0 \quad (3)$$

$$e(ta_j) = \sum_{i=0}^{j-1} \min \{ E_{max}, e(te_i) + e_{ij} \} * x_{ij} \quad (4)$$

$$d(ta_j) = \sum_{i=0}^{j-1} d(te_i) * x_{ij} \quad (5)$$

$$e(te_j) = e(ta_j) - e_j * x_j \geq E_{min} \quad (6)$$

$$d(te_j) = \min \{ D_{max}, d(ta_j) - d_j * x_j \} \geq 0 \quad (7)$$

$$(ta_j - te_i) * x_{ij} \geq T_{sij} \quad (8)$$

上述模型中,优化目标式(1)是使规划安排的观测任务的累计收益值最大化。约束条件中,式(2)和式(3)定义了初始时刻星载电池剩余的能量和星载存储器可用的存储空间;式(4)给定了任务开始成像时可用能量的计算方式,即前一个任务执行结束后的能量与到当前任务开始成像这段时间间隔

内能量变化量的和,同时限定了能量状态不能低于星载电池电量下限;式(5)给定了任务开始成像时剩余存储空间的计算方式,即为前一个任务执行结束后的存储空间状态;式(6)给定了任务成像结束时星载电池能量状态的计算方式,要求该时刻剩余的能量状态不能低于星载电池能量下限;式(7)给定了任务成像结束时星载存储器剩余存储空间的计算方式,要求该时刻剩余存储空间不得为负,避免存储数据溢出;式(8)要求两个连续执行的任务之间的时间间隔不能小于最短切换时间间隔。

4 序贯决策框架

为提高 AEOS 应对天上任务、资源以及外部环境动态变化的快速响应能力,本文采用类似于文献[16]的序贯决策框架来对 SOOTP 问题进行实时求解,如图 2 所示。其基本思路是:采用任务决策器按照时间先后顺序依次对每个观测任务进行决策,决策其是否应该被执行,直至最后一个任务。当任务决策器的输出结果为“True”时,表示当前任务将会被执行否则将会被舍弃。

由于观测任务之间在能量、存储等资源使用上存在激烈的竞争关系,因此,一个观测任务是否会被执行取决于卫星能量和存储等资源状态、观测任务自身属性信息以及其他观测任务的属性信息和分布情况等多重因素^[16]。为此,文献[16]中的任务决策器在决策每个观测任务的执行状态时,为保证决策的准确率,至少需要输入未来 3 个轨道圈次内观测任务的属性及分布等特征信息。

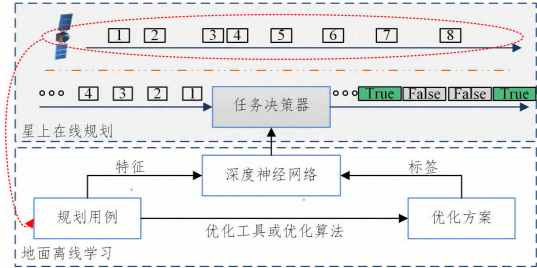


图 2 卫星观测任务序贯决策框架

Fig. 2 Framework of sequential decision-making for satellite observation tasks

可以看出,当任务决策器对多个连续相邻的观测任务进行决策时,它们的输入序列中包含了大量相同任务的特征信息。若对连续相邻的观测任务进行决策时能够重复使用这些特征信息,则可以大幅缩短任务决策时间,有效提升决策效率。基于上述思路,本文设计了新的观测任务序贯决策流程和决策模型,决策流程如算法 1 所示。

算法 1 观测任务序贯决策流程

输入:观测任务集合 $M = \{M_1, M_2, \dots, M_O\}$

输出:观测方案 selectedTask

1. selectedTask ← null
2. for $i = 1, 2, \dots, O$ // O 为规划时段内卫星绕地球飞行的圈数 * /
3. $of_{i,i+1,i+2} \leftarrow \text{getOrbitFeature}(M_i, M_{i+1}, M_{i+2})$ // 计算 3 个轨道圈次内所有任务的特征 * /
4. $encoded_of_{i,i+1,i+2} \leftarrow \text{orbitEncoder}(of_{i,i+1,i+2})$ // 计算 3 个轨道圈次内任务的抽象特征 * /


```

5.  for j=1,2,...,|Mij|/*|Mij|是当前轨道内任务数目*/
6.  if constraintsCheck(selectedTask,mij) then /*如果观测任务
    与已执行的任务冲突,返回 True,否则返回 False*/
7.      continue
8.  end if
9.  mfij←getTaskFeature(mij)/ *计算待决策任务特征*/
10. do←taskDecoder(encoded_ofi,i+1,i+2,mfij)/ *决策任务 mij的
    规划结果*/
11. if do is True then
12.     selectedTask←selectedTask∪mij
13. end if
14. end for

```

算法 1 的基本思路如下:当卫星从第 $i-1$ 个轨道圈次进行第 i 个轨道圈次时, $orbitEncoder()$ 函数批量计算出 $i, i+1, i+2$ 这 3 个轨道圈次(根据文献[16]中的实验结果确定而来)内所有任务的高度抽象特征信息(见算法 1 中的步骤 4),而对于第 i 个轨道圈次内的观测任务,则采用 $taskDecoder()$ 函数逐个计算是否应该被安排(见算法 1 中的步骤 10)。上述流程滚动进行,直到最后一个任务。与文献[16]中对每个任务的安排状态进行决策时都需要 $orbitEncoder()$ 和 $taskDecoder()$ 计算不同,算法 1 在每个轨道圈次只需进行一次 $orbitEncoder()$ 计算,从而可以有效提升决策效率。

$orbitEncoder()$ 和 $taskDecoder()$ 是图 2 所示任务决策器的核心,它们的具体实现细节将在下一节进行阐述。

5 基于注意力神经网络的观测任务决策模型

5.1 注意力神经网络模型及输入特征表示

基于上文建立的观测任务序贯决策框架,在对某个观测任务的调度结果进行决策时(下文称其为待决策观测任务),其他观测任务因收益值、观测开始时间、观测结束时间等属性信息不同,对待决策观测任务的影响程度也各不相同。为了获取对待决策观测任务影响较大的观测任务特征信息以及忽略不相关的信息,本文引入了注意力机制,并将其与循环神经网络相结合来捕获待决策观测任务与其他任务之间的依赖关系,实现对调度结果的准确决策。

本文设计了基于注意力神经网络的观测任务决策模型(Observation Task Decision-making Model Based on Attention Neural Network, AttNN)来构建任务决策器,该模型由编码网络和解码网络两部分组成,如图 3 所示。

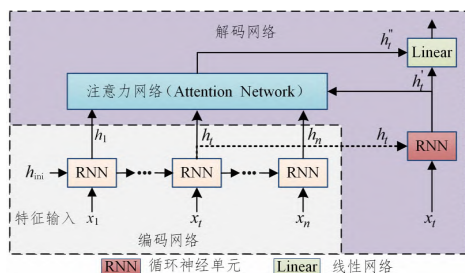


图 3 基于注意力神经网络的观测任务决策模型(电子版为彩图)

Fig. 3 Observation task decision-making model based on attention neural network

编码网络(图 3 中的灰色区域)是算法 1 中 $orbitEncoder()$

函数的具体实现,负责集中对每个轨道圈次内所有观测任务的属性及分布信息进行特征抽取,计算得到轨道圈次内所有任务的高度抽象特征信息;解码网络(图 3 中的紫色区域)是算法 1 中 $taskDecoder()$ 函数的具体实现,负责根据待决策观测任务的特征信息和编码网络输出的高度抽象特征信息对待决策任务的执行状态进行决策。

设编码网络的输入任务序列为 $M_{input} = m_1, m_2, \dots, m_n$ 。其输入特征表示为 $X_{input} = x_1, x_2, \dots, x_n$, 经过编码网络计算后可得到隐藏层输出 $H_{out} = h_1, h_2, \dots, h_n$, 其中 x_t 和 h_t 分别是第 t 个观测任务 m_t 的特征表示和隐藏层输出状态。

当决策是否安排第 t 个观测任务 m_t 时,首先,将其在编码网络对应的隐藏层状态 h_t 和特征表示 x_t 输入到解码网络的循环神经单元(图 3 中的红色方块),得到隐藏层输出状态 h'_t ;然后将 h'_t 和 $H_{out} = h_1, h_2, \dots, h_n$ 输入到注意力神经网络模块^[15](计算式见式(9)一式(11),得到其他任务与任务 m_t 相关的抽象特征信息 h''_t ;最后,将 $[h'_t, h''_t]$ 输入到线性网络中,经过 softmax 函数计算后便可得到最终决策结果。

$$u'_j = v^T \tanh(W_1 h_j + W_2 h'_t) \quad (9)$$

$$a'_j = \text{softmax}(u'_j) \quad (10)$$

$$h''_t = \sum_{j=1}^n a'_j h_j \quad (11)$$

式(9)一式(11)中, v, W_1 和 W_2 是注意力网络模块需要学习的参数,具体细节可参考文献[17]。

由于观测任务决策模型网络结构的改变,文献[16]中原有的部分任务特征信息无法通过计算得到,因此本文重新设计了观测任务输入特征表示,构建了包含任务属性、资源变化和任务冲突等信息在内的 11 个特征,如表 2 所列。

表 2 观测任务特征表示

Table 2 Feature representation of observation task

符号	含义
$te_i - ta_i$	观测任务 m_i 的成像时长
b_i	观测任务 m_i 的收益值
e_i	执行观测任务 m_i 需要消耗的能量
d_i	执行观测任务 m_i 需要占用的存储空间
$ta_i - te_{i-1}$	与前一任务之间的时间间隔
$e_{(i-1)i}$	$i=0$ 时等于当前星上电池的能量状态,否则为与前一观测任务的时间间隔内能量的变化量
$d_{(i-1)i}$	$i=0$ 时等于当前星载存储器的剩余的存储空间,否则为 0
$n_{pre,i}$	前序冲突任务数量,即成像开始时间大于 ta_i 且与任务 m_i 冲突的任务数量
$w_{pre,i}$	前序冲突任务平均收益值,即成像开始时间大于 ta_i 且与任务 m_i 冲突的任务的平均收益值
$n_{aft,i}$	后序冲突任务数量,即成像开始时间小于 ta_i 且与任务 m_i 冲突的任务数量
$w_{aft,i}$	后序冲突任务平均收益值,即成像开始时间小于 ta_i 且与任务 m_i 冲突的任务的平均收益值

5.2 决策模型训练方法

观测任务决策模型的训练过程分为两个阶段进行。

第一阶段,训练编码网络。本文采用 seq2seq(Sequence to Sequence)模型^[18]对编码网络进行训练。模型输入为任意连续 3 个轨道圈次内观测任务的特征表示,标签值是每个任务的规划结果,采用监督学习方法进行训练。训练结束后便可得到编码网络。

第二阶段,训练解码网络。如图 3 所示,训练解码网络时,模型输入为包括编码网络隐藏层输出和待决策任务的特征表示,标签值是任务的规划结果(0 或 1),采用监督学习方法进行训练即可。

在上述两个阶段中,编码网络和解码网络均采用交叉熵损失函数(见式(12))和 Adam 优化算法进行训练。

$$L = - \sum_{X \in D} \sum_{y=1}^K p(y, X) \log P(y|X) \quad (12)$$

其中, $p(y, X)$ 是训练样本 X 的独热编码分布函数(Gold One-hot Distribution), $P(y|X)$ 是各类的概率分布函数。

6 实验结果与分析

6.1 实验设置

本文采用文献[16]中的区域城市数据集和评价指标来验证所提方法的有效性。

(1) 实验数据

区域城市数据集包含 5 组规划场景,每组规划场景包含 120 个规划用例。5 组场景的规划用例的观测任务数量分别为 100, 200, 300, 400, 500。数据集中所有目标均位于北纬 $0^\circ \sim 60^\circ$ 、东经 $0^\circ \sim 150^\circ$ 的区域范围内。

(2) 评价指标

1) 平均收益(Average Profit),即单个场景下所有测试用例收益的平均值。数值越大,表明测试算法的优化性能越好。

2) 平均收益误差(Profit Gap),即单个场景下测试算法规划结果和基准算法规划结果的收益差与基准算法规划结果收益值比值的平均值。数值越小,表明测试算法的优化性能越好。

3) 平均响应时间(Response Time),即任务决策器/搜索算法对单个任务进行决策/规划所需要的平均计算时间,数值越小,表明算法的响应速度越快。

上述指标的具体计算式见文献[16]。

6.2 实验对比

为验证本文方法的有效性,本文以优化工具 ILOG CPLEX 的求解结果为参考基准,对比了文献[16]提出的基于长短期记忆网络(Long Short-Term Memory, LSTM)和基于门控循环单元(Gated Recurrent Unit, GRU)的两种网络决策模型(GL-LSTM 和 GL-GRU)、文献[10]提出的收益优先(Profit-based First, PF)启发式搜索算法和动态收益系统效用(Dynamic Profit System Benefit, d-PSB)启发式搜索算法。

本文算法 AttNN 的参数设置如下:编码网络采用双向 GRU 循环神经网络,隐藏层节点数为 32;解码网络循环神经单元为 GRU,隐藏层节点数为 64,线性模块为单层线性网络(64×2)。模型训练时学习率、批处理数目及迭代次数分别为 0.0001, 128, 50。对比算法均采用相应文献中设置的最优参数。

上述所有算法均基于 python 3.8.5 和 Pytorch 1.7.0 实现,运行在装有 Ubuntu 20.04 LTS 操作系统、配置为 i7-8550U 2.00GHz 处理器和 16GB 内存的计算机上。

表 3 列出了 AttNN, GL-LSTM, GL-GRU, PF, d-PSB 这 5 种算法的平均收益和收益误差。表格中数据的含义如下:以第 3 行第 2 列数据“1492(2.61%)”为例,“1492”“2.61%”分别表示 AttNN 算法的平均收益和收益误差。表 4 列出了上述 5 种算法的平均响应时间。

表 3 5 种算法的平均收益和平均收益误差

Table 3 Average profit and profit gap of five algorithms

任务数量	对比基准 CPLEX	对比算法				
		AttNN	GL-GRU	GL-LSTM	d-PSB	PF
100	1534	1492(2.61%)	1500(2.24%)	1502(2.19%)	1445(5.77%)	1508(1.66%)
200	2190	2114(3.57%)	2126(2.94%)	2124(3.05%)	2011(8.19%)	1976(9.78%)
300	2578	2487(3.68%)	2505(2.81%)	2504(2.86%)	2359(8.47%)	2148(16.65%)
400	2849	2730(4.28%)	2758(3.20%)	2756(3.29%)	2583(9.33%)	2235(21.56%)
500	3069	2933(4.55%)	2948(3.96%)	2949(3.95%)	2778(9.50%)	2277(25.83%)

表 3 中,任务数量为 100 时,PF 在规划收益上要高于其他 4 种算法,这是因为当参与规划的观测任务数目较少时,任务之间的冲突程度较小,选择收益值高的观测任务能够取得较好的效果。随着观测任务数目的增多,任务之间的冲突程度增大,AttNN, GL-LSTM 和 GL-GRU 这 3 种深度学习方法在规划收益上远高于 PF 和 d-PSB 这两种启发式搜索算法,这是因为学习算法通过对历史任务规划场景的学习,理解了观测任务决策的相关知识,可以直接指导观测任务规划,这说明了机器学习算法解决观测任务规划问题的可行性与有效性。将 AttNN 与 GL-LSTM 和 GL-GRU 进行对比,虽然 AttNN 的收益误差高于 GL-LSTM 和 GL-GRU,但差距不超过 1%,整体误差控制在 4.6% 以内。

表 4 5 种算法的平均响应时间

Table 4 Response time of five algorithms

(单位:ms)

任务数量	对比算法				
	AttNN	GL-GRU	GL-LSTM	d-PSB	PF
100	0.52	3.0	3.0	0.22	0.21
200	0.52	3.0	3.1	0.33	0.26
300	0.53	3.1	3.1	0.38	0.28
400	0.53	3.2	3.2	0.42	0.29
500	0.54	3.2	3.2	0.45	0.29

表 4 中,AttNN, GL-LSTM 和 GL-GRU 的平均响应时间要长于 PF 和 d-PSB,这是因为 PF 和 d-PSB 只需计算各个任务的排序指标即可,而 3 种深度学习方法需要各个任务的高度抽象特征,耗时更长。值得注意的是,AttNN 的平均响应时间远短于 GL-LSTM 和 GL-GRU,这是因为 AttNN 算法中编码网络在每个轨道圈次只需计算一次,GL-LSTM 和 GL-

GRU 在每次任务决策时编码网络均需要计算,而编码网络又是任务决策器中最耗时的计算部分,因此 AttNN 可以节省大量的计算时间。

同时可以看到,AttNN 和 PF 以及 d-PSB 的平均响应时间处于同一量级,这是因为 AttNN 是按照时间轴的顺序依次对任务进行决策,可以避免过多的约束检测操作,其响应时间主要取决于模型的计算时间。而 PF 和 d-PSB 算法是根据观测任务的排序依次将任务插入到观测方案中,每次插入新观测任务时,都要对当前任务后的所有观测任务进行任务时序、能源存储等约束检测,以保证观测方案的可行性,而约束检测过程会耗费大量的计算时间。观测任务动态达到时,PF 和 d-PSB 算法的响应时间很大程度上取决于动态观测任务的数量规模,数目越多,响应时间就越长。因此,在大量观测需求到达的情况下,AttNN 将体现出更强的响应能力,这是序贯决策框架的天然优势。

综合表 3 和表 4 的结果可以看出,AttNN 的响应时间不及 GL-LSTM 和 GL-GRU 的 1/5,与 PF 和 d-PSB 处于同一量级;AttNN 的收益误差与 GL-LSTM 和 GL-GRU 接近,远大于 PF 和 d-PSB 启发式搜索算法。可见,本文算法 AttNN 在响应时间和平均收益上更加均衡,适应性更好。

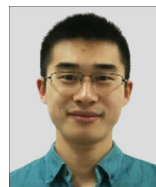
上述实验结果及分析验证了本文提出的基于注意力神经网络的观测任务序贯决策算法的可行性与有效性。

结束语 本文对对地观测卫星上自主任务规划问题进行了研究,改进了观测任务序贯决策流程,设计了基于注意力神经网络的观测决策模型、任务特征表示方法以及模型训练方法,提出了一种基于注意力神经网络的观测任务序贯决策方法。实验结果表明,该方法可以有效地解决卫星上自主任务规划问题,且具有更高的决策效率。

未来将主要集中于敏捷卫星观测任务序贯决策方法的研究。敏捷卫星具有更高的观测灵活度,需要同时规划观测任务的观测顺序和观测时间。如何将深度神经网络应用于敏捷卫星任务规划,值得我们深入研究。

参 考 文 献

- [1] DU Y, WANG T, XIN B, et al. A data-driven parallel scheduling approach for multiple agile earth observation satellites[J]. IEEE Transactions on Evolutionary Computation, 2020, 24(4): 679-693.
- [2] DENG M, LIU B, LI S, et al. A two-phase coordinated planning approach for heterogeneous earth-observation resources to monitor area targets[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 99: 1-16.
- [3] LIU M, ZHANG Q. Remote Sensing Data Transmission Mechanism Based on Distributed Cluster Architecture[J]. Computer Engineering, 2021, 47(10): 180-185.
- [4] XIANG S, CHEN Y G, LI G L, et al. Review on satellite autonomous and collaborative task scheduling planning[J]. Acta Automatica Sinica, 2019, 45(2): 252-264.
- [5] CHIEN S, SHERWOOD R, BURL M, et al. A demonstration of robust planning and scheduling in the techsat-21 autonomous spacecraft constellation [J]. Ear Nose & Throat Journal, 2014, 86(8): 506-511.
- [6] CHIEN S, RABIDEAU G, TRAN D, et al. Scheduling science campaigns for the rosetta mission: a preliminary report [C]// International Workshop on Planning and Scheduling for Space, 2013: 1-8.
- [7] BEAUMET G, VERFAILLIE G, CHARMEAU M. Feasibility of autonomous decision making on board an agile earth-observing satellite [J]. Computational Intelligence, 2011, 27(1): 123-139.
- [8] LIU S, CHEN Y, XING L, et al. Time-dependent autonomous task planning of agile imaging satellites [J]. Journal of Intelligent & Fuzzy Systems, 2016, 31(3): 1365-1375.
- [9] CHU X, CHEN Y, TAN Y. An anytime branch and bound algorithm for agile earth observation satellite onboard scheduling [J]. Advances in Space Research, 2017, 60(9): 2077-2090.
- [10] LI G, XING L, CHEN Y. A hybrid online scheduling mechanism with revision and progressive techniques for autonomous Earth observation satellite [J]. Acta Astronautica, 2017, 140(1): 308-321.
- [11] SU J M, LIU H F, XIANG F T, et al. Survey of Interpretation Methods for Deep Neural Networks [J]. Computer Engineering, 2020, 46(9): 1-15.
- [12] WANG H J, YANG Z, ZHOU W G, et al. Online scheduling of image satellites based on neural networks and deep reinforcement learning [J]. Chinese Journal of Aeronautics, 2019, 32(4): 1011-1019.
- [13] LI C, CAUSMAECKER P D, CHEN Y W. Data-driven onboard scheduling for an autonomous observation satellite [C]// Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018: 5773-5774.
- [14] LU J, CHEN Y, HE R. A learning-based approach for agile satellite onboard scheduling [J]. IEEE Access, 2020, 99(8): 16941-16952.
- [15] ZHAO X, WANG Z, ZHENG G. Two-phase neural combinatorial optimization with reinforcement learning for agile satellite scheduling [J]. Journal of Aerospace Information Systems, 2020, 17(7): 1-12.
- [16] PENG S, CHEN H, DU C, et al. Onboard observation task planning for an autonomous earth observation satellite using long short-term memory [J]. IEEE Access, 2018, 6(1): 65118-65129.
- [17] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [C]// International Conference on Learning Representations, 2015: 1-15.
- [18] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks [C]// 27th International Conference on Neural Information Processing Systems, 2014: 3104-3112.



PENG Shuang, born in 1990, Ph.D, lecturer, is a member of China Computer Federation. His main research interests include satellite intelligent scheduling and machine learning.



CHEN Hao, born in 1982, Ph.D, professor, master supervisor, is a member of China Computer Federation. His main research interests include computational intelligence, machine learning and satellite intelligent scheduling.

(责任编辑:杨雪敏)