

# 基于近似动态规划的动态车辆调度算法

李 雪 聂兰顺 齐文艳 战德臣

哈尔滨工业大学, 哈尔滨, 150001

**摘要:**针对物流配送服务业中, 车辆调度问题日渐呈现任务规模大, 车辆类型多、属性多, 调度实时性要求越来越高等特点, 提出了基于近似动态规划的动态车辆调度算法。根据当前的任务需求与车辆状态以及相应的约束条件作出相应的调度, 并且对一些样本进行训练, 得到了一个近似价值函数。通过该价值函数, 即可对任务迅速作出相应的决策。仿真模拟实验证明了该算法的有效性和优越性。

**关键词:**服务资源; 近似动态规划; 动态调度; 价值函数

**中图分类号:** TP311.52

**DOI:** 10.3969/j.issn.1004-132X.2015.05.020

## An Algorithm of Dynamic Vehicle Scheduling Problem Based on Approximate Dynamic Programming

Li Xue Nie Lanshun Qi Wenyan Zhan Dechen

Harbin Institute of Technology, Harbin, 150001

**Abstract:** Vehicle scheduling in service industry of logistics distribution was presenting features including the tasks tended to be of large scale, vehicles were multi-type and had multiple attributes as well as high demands for real-time scheduling. To solve these problems, this paper proposed a dynamic vehicle scheduling algorithm based on the approximate dynamic programming. An approximate value function was obtained through training of some samples, and according to mission requirements, vehicle state and conditions, and quick scheduling decisions could be made with the value function. The simulation test has proved the correctness and effectiveness of the algorithm.

**Key words:** service resource; approximate dynamic programming; dynamic scheduling; value function

## 0 引言

针对车辆动态调度问题, 国内外专家学者开展了一系列研究。Gendreau 等<sup>[1]</sup>着重研究了车辆动态调度问题中出现的各种不确定性信息的影响, 指出在求解此问题时, 对这些不确定性信息应加以全面考虑; Powell<sup>[2]</sup>详细分析了车辆动态调度问题中一类随机车辆调度模型, 采用改进的 A-priori 两阶段优化方法求解了该问题; Minkoff<sup>[3]</sup>以马尔可夫决策模型为基础, 完成了车辆动态调度问题基于马尔可夫决策过程的建模求解, 研究提出的算法在中小规模(10 个需求)的车辆动态调度问题求解中可以得到比较满意的解, 但因其模型的局限性, 算法对大规模的问题难以求解。针对动态车辆调度问题实时性强的特点, 张景玲等<sup>[4]</sup>、王旭等<sup>[5]</sup>研究了车辆的动态调度问题, 通过基于两阶段优化的方法对该类问题进行了有效求解。袁建清<sup>[6]</sup>以解决车辆利用效率最大化为目标, 建立了几类随机数学模型, 提出了相应的智能算法, 解决了车辆调度的不确定调度问题。

文献[7-9]针对车辆动态调度中的不同问题提出了相应的解决方法。

本文以军事行动中车辆动态调度问题为背景, 提出了基于近似动态规划的车辆动态调度算法。通过对车辆调度问题进行形式化描述, 利用近似动态规划方法对车辆动态调度问题进行建模, 根据近似动态规划的思想, 设计实现求解大规模、多类型车辆调度的算法, 并对算法进行了仿真性实验, 验证了算法的有效性和优越性。

## 1 动态车辆调度的问题

车辆调度问题对实时性有较高要求, 即在尽可能短的时间内, 通过合理的运输方式、运输路径、运输工具组合来完成调度任务, 是动态车辆调度问题领域关注的重点。对于动态车辆调度问题, 本文以一个有关的军事任务中的车辆调度情景予以描述, 如图 1 所示。

在某次军事演习中, 共涉及有 1 个车辆调度中心和  $N$  个驻防要点, 车辆调度中心拥有载重车、乘坐车、牵引车和特种车四种类型的运力资源, 共  $K$  辆运输工具。每个驻防要点拥有兵员、物资、装备等参演要素。演习中, 需要在这  $N$  个驻防要点之间完成兵员、物资和装备的调运服务。演习过程中无法预知哪个驻防要点具有运输任务

收稿日期: 2013-04-23

基金项目: 国家自然科学基金资助项目(61273038); 国家科技支撑计划资助项目(2013BAH17F03); 国家重点基础研究发展计划(973 计划)资助项目(2010CB328004); 山东省自主创新成果转化重大专项(2012ZH1A0411)

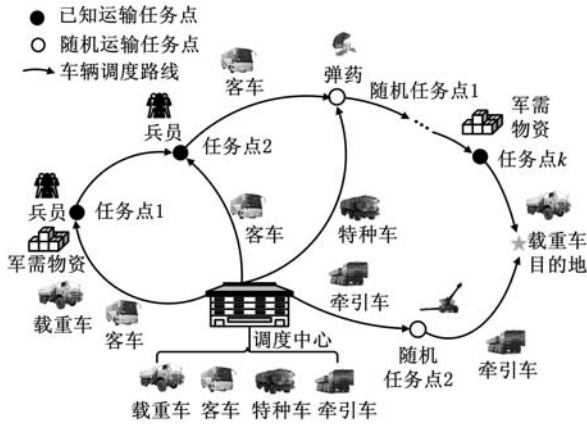


图 1 军演中车辆调度情景图

请求。根据描述情况,可对上述场景进行抽象,得到如表 1 所示的信息。

表 1 大规模、多类型动态车辆调度问题范围约束

| 问题要素     | 描述   |
|----------|--|
| 车辆种类和数量  | 多类型,多辆   |
| 车辆载重/员约束 | 不同类型车辆具有不同载重/员,相同类型车辆具有相同载重/员                  |
| 运输对象     | 多类型,例如物资、兵员、弹药等                                |
| 运输任务     | 完成各任务点之间运输对象的调运,任务数量为大规模                       |
| 时间约束     | 执行任务的车辆必须在任务起始时间前到达任务点,在任务结束时间前完成任务,否则任务失败     |
| 不确定性因素   | 每个时刻可能会有随机出现的任务,这样的任务信息是不确定的;执行任务的车辆出现故障具有不确定性 |

将本文研究的问题看作一个系统,问题中每个时刻的调度场景就可看作是该系统的一个状态,那么每个时刻的系统状态与该时刻的调度决策是一一对应的,不同的调度决策导致系统到达不同的状态,因此,每个时刻不同的系统状态价值,可以反映不同调度决策的优劣。由此,文中系统状态价值的含义可以描述为:每个时刻不同的调度决策会对系统的现状和将来产生不同的贡献,由此,每个时刻的系统状态价值就是该时刻对应调度决策对系统的贡献值。

本文涉及的大规模、多类型车辆动态调度问题中,由于不同的运输车辆具有不同的代价权重,因此,当执行不同任务时,动用的车辆不同,相应的收益权重也不同。例如执行任务类别  $m_i$ , 它的回报权重为  $\zeta_i$ ,  $t$  时刻需要被满足的任务数(决策作用前)为  $n_{it}$ ,  $t$  时刻满足的任务数(决策作用后)为  $n'_{it}$ ; 车辆类别为  $c_j$ , 代价权重为  $\eta_j$ ,  $t$  时刻可调度的车辆数(决策作用前)为  $k_{jt}$ ,  $t$  时刻可调度的车辆数(决策作用后)为  $k'_{jt}$ , 那么定义贡献函数为

$$\sum_{i=1}^n n'_{it} \zeta_i / \sum_{j=1}^m k_{jt} \eta_j。$$

由此,我们用系统状态价值来定义本文研究问题的优化目标:根据每个时刻不断出现的运输任务请求和不断变化更新的车辆状态,在一定条件下(如运输任务类型、任务起止时间、车辆剩余载重、单次最大行驶里程等),动态查询所有车辆状态,挑选合适的车辆,规划合理的路线,尽可能地满足任务点的运输任务请求,使得每个时刻的系统状态价值最大。

## 2 动态车辆调度问题模型

根据上述情景分析,对此调度场景建立相应的模型,主要包括车辆资源、任务信息和调度决策等。

### 2.1 车辆资源

车辆资源建模的基本思想是抽象出车辆资源的重要属性,明确车辆资源的使用规则,从而约束车辆属性向量的空间取值。车辆属性主要包括静态属性和动态属性:静态属性描述车辆资源的基本特点;动态属性刻画车辆资源的状态。

$C$  为车辆资源类型集合,  $C = \{c_1(\text{载重车}), c_2(\text{客车}), c_3(\text{特种车}), c_4(\text{牵引车})\}$ ;  $\mathbf{a}$  为车辆资源属性向量,  $\mathbf{a} = (a_1(\text{车辆编号}), a_2(\text{车辆类型}), a_3(\text{实时位置}), a_4(\text{额定载重/员}), a_5(\text{剩余载重/员}), a_6(\text{额定里程}), a_7(\text{剩余行驶里程}), a_8(\text{平均速度}), a_9(\text{下一任务点}), a_{10}(\text{起效时间}), a_{11}(\text{上一次接受调度的任务点}))$ ;  $\mathbf{A}$  为车辆资源的向量空间;  $R_{t',t'a}$  为  $t'$  时刻获知的,具有属性  $\mathbf{a}$ 、在  $t'$  时刻可用的车辆资源的数量,  $t' > t$ ;  $R_{ta}$  为  $t$  时刻获知的,具有属性  $\mathbf{a}$ 、在当前时刻可用的车辆资源的数量,  $R_{ta} = (R_{t',t'a})_{t'=t}$ ;  $\hat{R}_{ta}$  为  $t$  时刻获知的,  $t-1$  时刻和  $t$  时刻之间发生的,由外部信息导致的、具有属性  $\mathbf{a}$  的、可调度车辆资源的数量,其中外部信息包括:车辆故障停驶、故障车辆恢复行驶等。

通过以上对车辆资源的建模,可以得到:  $t$  时刻,可以调度的具有属性  $\mathbf{a}$  的车辆资源的数量为

$$R_{ta} = R_{t-1,ta} + \hat{R}_{ta}$$

$t$  时刻,可以调度的车辆资源的数量为

$$R_t = (R_{ta})_{a \in A}$$

### 2.2 任务信息

任务请求信息也可以看作是系统资源,为了刻画运输任务的多方面属性以及运输任务的静态属性和动态属性,笔者建立了调度决策模型。通过属性向量来描述和刻画运输任务的状态;通过明确运输任务的使用规则,来约束其属性向量的空间取值。

$M$  为运输任务类型集合,  $M = \{m_1(\text{物资运$

输),  $m_2$ (兵员运输),  $m_3$ (弹药运输),  $m_4$ (装备运输)};  $\mathbf{b}$  为运输任务属性向量,  $\mathbf{b} = \{b_1$ (任务编号),  $b_2$ (任务类型),  $b_3$ (任务开始时间),  $b_4$ (任务结束时间),  $b_5$ (任务起始点),  $b_6$ (任务结束点),  $b_7$ (任务量),  $b_8$ (任务完成状态)};  $\mathbf{B}$  为运输任务属性的空间向量;  $M_{t,i,b}$  为  $t$  时刻获知的, 在  $t'$  时刻需要被满足的、具有属性  $b$  的运输任务请求的数量;  $M_{tb}$  为  $t$  时刻获知的, 在当前时刻需要被满足的、具有属性  $b$  的运输任务请求的数量, 即  $M_{tb} = (M_{t,i,b})_{i=t}$ ;  $\bar{M}_{tb}$  为  $t$  时刻获知的, 在  $t-1$  时刻和  $t$  时刻之间随机出现的, 具有属性  $b$  的运输任务请求的数量。

那么,  $t$  时刻需要被满足的、具有属性  $b$  的运输任务请求的数量为

$$M_{tb} = M_{t-1,b} + \bar{M}_{tb}$$

$t$  时刻需要被满足的任务数量为

$$M_t = (M_{tb})_{b \in B}$$

### 2.3 调度决策

调度决策属于动态系统的内部信息, 为了刻画调度决策的内容以及调度决策如何起效作用于车辆资源和运输任务, 对调度决策的建模要从对车辆资源和运输任务的状态影响出发, 抽象其重要属性, 通过定义调度决策的策略集, 来约束其属性向量的空间取值。

$\mathbf{d}$  为调度决策的属性向量,  $\mathbf{d} = (d_1$ (当前派遣),  $d_2$ (暂不派遣),  $d_3$ (执行车辆编号),  $d_4$ (执行任务编号),  $d_5$ (派遣时刻));  $\mathbf{D}_a$  为可以作用于具有属性  $\mathbf{a}$  的车辆资源向量;  $\Pi$  为可行调度策略的集合。调度策略是指在给定系统状态信息的前提下, 决定一个调度决策的规则。在本文的研究中, 调度策略由贡献函数(反映当前调度决策对系统当前贡献的影响)和近似价值函数(反映当前调度决策对系统未来的影响)共同来反映。  $x_{tad}$  为  $t$  时刻, 具有属性  $\mathbf{a}$  的, 被决策  $\mathbf{d}$  作用的车辆的数量, 则

$$\sum_{\mathbf{d} \in \mathbf{D}_a} x_{tad} = R_{ta}, \forall \mathbf{a} \in \mathbf{A} \quad (1)$$

$$x_{tad} \geq 0, \forall \mathbf{a} \in \mathbf{A}, \mathbf{d} \in \mathbf{D}_a$$

$\sigma_{tad}$  为决策结果指示函数, 用来捕获决策的结果, 且

$$\sigma_{tad} = \begin{cases} 1 & \mathbf{d} = d_1 \\ 0 & \text{其他} \end{cases}$$

那么  $t$  时刻, 被派遣执行运输任务的、具有属性  $\mathbf{a}$  的车辆数量为  $\sigma_{tad} x_{tad}$ ;  $\chi_t$  为  $t$  时刻, 在给定有效信息下的可行调度策略的集合。

为了通过数学形式来反映决策结果, 需要定义一个决策函数, 一些调度策略, 在每个取样时

刻, 给定系统的状态信息, 返回调度决策。

$X_t^\pi(R_t)$  为决策函数。  $t$  时刻, 在调度策略  $\pi$  下, 给定车辆资源状态  $R_t$ , 返回一个决策值  $x_t$ , 其中  $R_t$  为  $t$  时刻系统的状态信息:

$$X_t^\pi(R_t) = \operatorname{argmax}\{C_t(R_{t-1}, \omega_t, x_t) + \gamma \bar{V}_t(R^M(R_{t-1}, \omega_t, x_t))\} \quad (2)$$

其中,  $\gamma$  为折扣因子, 介于 0 和 1 之间, 是指价值经过一段时间后等同的现在的价值, 因为价值函数是对未来的一个预测价值, 所以需要加上一个折扣因子;  $\bar{V}_t$  为近似价值函数;  $C_t(R_{t-1}, \omega_t, x_t)$  为系统贡献函数;  $R^M(R_{t-1}, \omega_t, x_t) = R_t$  为车辆资源状态转换函数;  $\omega_t$  为  $t$  时刻系统外部信息的一次取样。

本文在车辆调度决策过程中, 考虑了两种车辆调度方式: 一是单车多任务, 二是多车单任务。

此外, 为了计算的方便, 对时间采取离散时间模型, 如图 2 所示。在前述的对车辆资源、运输任务、调度决策的符号中, 右下角的时间角标“ $t$ ”, 表示的是离散时间点  $t$  时刻或第  $t$  期, 第  $t$  期指  $t = (t-1, t]$ 。

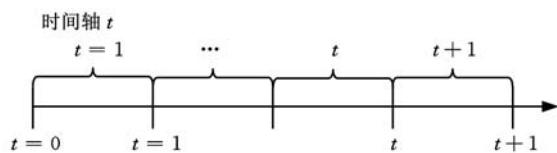


图 2 取样时刻模型

### 2.4 目标方程

把大规模、多类型车辆动态调度问题看作是一个“动态系统”, 把每次作调度决策的场景看作是该系统在时间轴上的一个“状态” $S_t$ 。  $S_t$  由车辆资源状态  $R_{t-1}$ 、运输任务信息  $M_t$  和调度决策  $x_t$  共同描述。  $S_t$  的价值是由贡献函数和近似价值函数共同决定。贡献函数捕获调度决策  $x_t$  对当前系统状态的影响; 近似价值函数捕获调度决策  $x_t$  对未来系统状态的影响。本文优化目标为: “每期决策, 使得在尽可能完成运输任务的前提下, 动用的车辆数最少; 长期目标是在完成尽可能多的任务前提下, 车辆动用率最低。”那么, 大规模、多类型车辆动态调度问题的目标方程可以形式化表达为

$$\hat{v}_t^n = \max_{x_t \in \chi_t} C_t(R_{t-1}^n, \omega_t^n, x_t) + \gamma \bar{V}_t^{n-1}(R^M(R_{t-1}^n, \omega_t^n, x_t)) \quad (3)$$

其中,  $x_t$  为式(3)的解,  $\hat{v}_t^n$  为系统状态从  $S_t$  条件转移到  $S_{t+1}$  的近似价值的数学期望。

当然, 式(3)还要满足一定的约束条件: ①调度决策作用的车辆资源数量不能超过当前已知的可调度车辆资源的数量; ②每个时刻的调度决策

满足的任务请求数不能超过当期已知的任务请求数；③调度决策作用的车辆资源数量、运输任务数量都是正整数。

### 3 基于近似动态规划的动态车辆调度方法

动态规划是基于多阶段决策过程寻优问题提出的,广泛应用于工程学、运筹学、经济学等多个领域<sup>[10]</sup>。但是,经典动态规划所面临的“维数灾难”使其只能解决小规模问题,限制了其应用<sup>[11]</sup>。通过上面的建模,本文采用近似动态规划(ADP)的思想设计动态车辆调度算法。

#### 3.1 基本思路和设计流程

基于 ADP 方法求解大规模、多类型的车辆动态调度问题需要划分为两个阶段,第一阶段是训练获取近似价值函数的表达式,第二阶段是应用训练得到的近似价值函数的表达式指导车辆调度。ADP 在训练数据阶段是用本次系统产生的数据去更新上一次假设的数据,即将来对过去的影响,不断地更新进而产生出近似价值函数;在应用阶段就是利用训练阶段产生的近似价值函数来生成任务到来时的决策,即对未来的影响。

在第一阶段中,算法基于“观察-更新”的思想,预先设定总的取样次数  $N$  和取样路径长度  $\tau$ , 给定每条取样路径上每个取样时刻对应的系统状态近似估计值的初值  $\bar{V}_i^0$ , 然后利用价值迭代的策略,迭代计算目标方程(式(3))。在每次迭代中,通过在取样路径  $\omega^n$  计算得到的新的系统状态值  $\hat{v}_i^n$  更新上一时刻系统状态的近似估计值  $\bar{V}_{i-1}^n$ , 以此来不断逼近系统状态的真实值。算法运行最后得到  $N$  组系统状态近似值  $\bar{V}_i^n = \{(\bar{V}_i^n)_{i=1}^{\tau}, (\bar{V}_i^n)_{i=1}^{\tau}, \dots, (\bar{V}_i^n)_{i=1}^{\tau}\}$ , 然后通过一组达到稳态的系统状态近似值,以线性拟合的思想,拟合这组稳态值,来得到近似价值函数的线性表达式。

因此,第一阶段算法的输入是仿真得到的任务信息,输出为训练周期中每期系统状态价值的近似值。通过仿真任务信息来获得、辨识和测量训练阶段算法的各种参数,比如折扣因子、步长以及系统状态初值等。在第二阶段,应用第一阶段训练得到的近似价值函数表达式,根据当前的运输任务信息,求解决策函数(式(2))以得到优化调度决策  $x_t$ 。因此,第二阶段算法的输入为当前运输任务信息,算法的输出为优化的调度决策  $x_t$ 。

#### 3.2 调度策略的启发式规则

在求解大规模、多类型车辆动态调度问题中,本文中车辆调度策略的启发式算法规则集如下:

(1) 对于每期出现的新运输任务,尽量从已经派出在外执行任务的车辆中挑选满足新运输任务要求的车辆,而尽量避免从调度中心增派车辆去满足新任务,以此来减少每期的车辆动用数量。

(2) 对于当期出现的多个运输任务,按照任务开始时间的紧急程度,优先满足任务开始时间早的任务。

(3) 对于在当期随机出现的运输任务请求,在任务开始时间和任务量满足的前提下,优先考虑与现有任务是否可以合并执行,以减少车辆动用的数量。

(4) 对于可以满足某一运输任务的多辆可调度车辆,先将这些车辆按照剩余载重的大小进行排序,然后依次挑选剩余载重大的车辆去执行该运输任务;对于剩余载重也相同的车辆,按照可以到达任务起始点的时间排序,依次选择可以最早到达任务起始点的车辆执行该运输任务,这样可以在多车单任务中,减少车辆动用的数量,从而降低车辆的动用率。

启发式规则的输入为当前时刻的运输任务信息,即需要被满足、具有某属性的多个运输任务;输出为可调度的车辆序列和已调度的车辆序列。算法具体步骤描述如下。

(1) 查询当期任务信息  $M_t$ , 按任务类型分类汇总得到每种类型任务数量  $M_{tb_2}$ 。

(2) 对于当期出现的每个任务  $M_{tb}$ , 按当期任务的开始时刻从小(早)到大(晚)排序。

(3) for 当期出现的、开始时间最早的任务:

do 按任务类型要求、开始时间要求查询是否有在外执行任务的、可调度的相应类型的车辆资源状态。

if 有在外执行任务的、可调度的车辆,  
do 返回在外执行任务的、可调度的车辆资源序列。

else 查询在调度中心的车辆资源,返回可调度的车辆资源序列;

$$R_{t,a_2=1,a_3 \neq 0} \quad R_{t,a_2=2,a_3 \neq 0} \quad R_{t,a_2=3,a_3 \neq 0} \quad R_{t,a_2=4,a_3 \neq 0}$$

(4) 将步骤(3)中得到的可调度车辆按剩余载重/员从大到小进行排序,得到每种类型可调度的车辆序列:

$$R_{t,a_2=1,a_5} \quad R_{t,a_2=2,a_5} \quad R_{t,a_2=3,a_5} \quad R_{t,a_2=4,a_5}$$

(5) 对步骤(4)中挑选出来的可调度车辆序列中,再对剩余载重相同的车辆按照起效时间从小(早)到大(晚)进行排序,得到每种类型可调度的车辆新序列如下:

$$R_{t,a_2=1,a_5,a_{10}} \quad R_{t,a_2=2,a_5,a_{10}} \quad R_{t,a_2=3,a_5,a_{10}} \quad R_{t,a_2=4,a_5,a_{10}}$$

(6) 计算单车是否可以满足该任务。

if 单车满足, do 转至步骤(7), else 转至步骤(8)。

(7) 从步骤(5)中挑选第一辆车。

(8) 从步骤(3)中依次挑选车辆,直到车辆组合剩余载重之和满足任务量要求。

(9) 按照贡献函数的定义式计算不同调度决策的贡献值,按当前最大贡献值对应的调度决策调度车辆执行任务。

(10) 将当期没有车辆满足的任务顺延至下一期转至步骤(1)。

### 3.3 采用价值迭代和平滑策略训练近似价值函数的算法设计

大规模、多类型车辆调度问题训练阶段的算法采用价值迭代和平滑策略来获取系统状态的真实值。具体算法步骤如下:

(1) 初始化。

① 为各个时刻的系统近似价值赋初值,  $\bar{V}_t^0 = 0, t \in \{1, 2, \dots, 30\}$ ;

② 设置  $n=1, N=100, n$  为取样路径标记,  $N$  为总的取样次数;

③ 初始化车辆资源状态  $R_0^1$ 。

(2) 选择一条取样路径  $\omega^n, \omega_t^n = M_t^n$ 。

(3) 对于训练阶段的每一个取样时刻,  $t=1, 2, \dots, 30$ 。

① 令  $\omega_t^n = W_t(\omega^n)$ , 进行取样的实现;

② 调用前述的启发式规则算法, 筛选得到最优的执行车辆;

③ 将执行车辆中可以推迟派遣的, 推迟一期派遣;

④ 根据步骤 ③ 计算式(3),  $x_t^n$  为式(3)的解。

⑤ 更新价值函数:

$$\bar{V}_{t-1}^n(R_{t-1}^n) = (1-\alpha)\bar{V}_{t-1}^{n-1}(R_{t-1}^n) + \alpha\hat{v}_t^n \quad (4)$$

⑥ 计算车辆资源状态转换函数, 更新车辆资源状态:

$$R_t^n = R^M(R_{t-1}^n, \omega_t^n, x_t^n)$$

(4)  $n$  加 1, 如果  $n \leq N$ , 跳转至步骤(2)。

(5) 返回每条取样路径的每组状态价值, 即  $(\bar{V}_t^n)_{t=1}^{10}$ 。

定义的近似价值函数是关于车辆资源状态和运输任务的线性函数, 而车辆资源状态和运输任务状态都是用向量描述的, 如果以向量进行近似价值函数计算, 其取值空间巨大, 容易陷入“维度灾难”, 难以求解; 因此, 笔者对这些向量的维度进

行了一定程度的聚集, 即忽略对车辆调度决策影响不大的维度。对于本文中的大规模、多类型车辆动态调度问题, 每期可调度的车辆数量和已满足的任务数量对于车辆调度决策影响较大, 因此, 将车辆资源属性向量  $(R_{t,a}^n)_{a \in A}$  聚集为当前时刻可调度的车辆数数量  $(R_{t,a}^n)_a$ ; 将运输任务属性向量聚集  $(M_{t,b}^n)_{b \in B}$  聚集为当前已被满足的运输任务数量  $(M_{t,b}^n)_b$ 。

接下来可计算近似价值函数的线性表达式:

$$\bar{V}_t^n = \theta_1 (R_{t,a}^n)_a + \theta_2 (M_{t,b}^n)_b + \theta_3$$

其中,  $\theta_1, \theta_2$  和  $\theta_3$  为待定参数, 根据上述 ADP 求解问题的算法步骤(5)达到稳态的一组有效值, 采用线性回归的方法求解得到待定参数  $\theta_1, \theta_2$  和  $\theta_3$ , 从而得到近似价值函数的线性表达式。

### 3.4 应用近似值函数进行大规模、多类型车辆调度算法设计

大规模、多类型车辆动态调度问题应用阶段算法是对训练阶段获得的近似价值函数进行调度应用, 具体算法步骤如下:

(1) 初始化车辆资源的状态  $R_0$ 。

(2) 输入当前时刻的运输任务信息  $M_t$ 。

(3) 调用前述的启发式规则算法, 求解决策函数:

$$X_t^r(R^{r-1}, \omega_t) = \operatorname{argmax}_{x_t \in X_t} C_t(R_{t-1}, \omega_t, x_t) + \gamma \bar{V}_t(R^M(R_{t-1}, \omega_t, x_t)) \quad (5)$$

其中, 调度决策  $x_t$  为式(5)的解。

(4) 更新车辆资源状态:

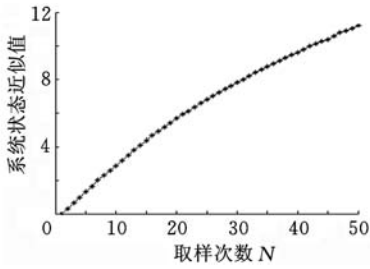
$$R_t = R^M(R_{t-1}, \omega_t, x_t)$$

## 4 实验

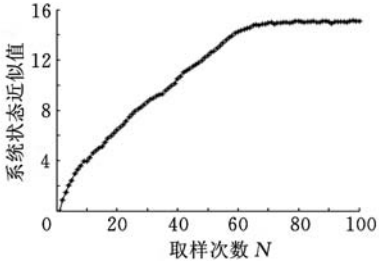
### 4.1 实验场景以及训练结果

根据上文中算法的描述, 进行了相应的实验。实验过程中, 假定有 4 种不同类型的车辆, 每种类型的车辆有 10 辆, 10 个任务点, 4 种不同的任务。价值迭代算法需要为其设计合理的收敛准则。实验中, 在取样时间轴上, 具有相同周期长度和固定期数的一组连续的系统状态, 本文尝试分别取样 50 次和取样 100 次, 观察每条取样路径上某一相同时刻系统状态价值的近似值是否趋于稳态, 用 MATLAB 分析, 结果如图 3 所示。

由图 3 观察比较可以发现: 算法在迭代 50 次后系统状态近似值依然呈现出稳步上升趋势, 说明值迭代策略还未逼近到系统状态的真实值; 在迭代 100 次后, 观察发现系统状态近似值已经趋于稳态, 说明值迭代策略已经逼近到系统状态的真实值, 算法已经收敛, 所以算法可以终止。



(a) 每条取样路径上  $t_1$  时刻的系统状态近似值(50次)



(b) 每条取样路径上  $t_1$  时刻的系统状态近似值(100次)

图3 算法迭代 50 次和 100 次后系统近似值变化图

取第 100 次迭代的最后一组系统状态近似值进行拟合求解,求得近似价值函数的线性表达式,如表 2 所示。

表 2 选取第 100 次取样的系统状态的近似值进行近似值函数拟合

| 当前可调度车辆数 $R$ | 当前已被满足任务数 $M$  | 当前时刻系统近似值 $V$ |
|--------------|----------------|---------------|
| 17           | 51             | 15.15615      |
| 17           | 47             | 15.1767       |
| 15           | 56             | 15.16908      |
| 13           | 69             | 15.20471      |
| 9            | 79             | 15.22799      |
| 15           | 56             | 15.18479      |
| 12           | 71             | 15.23271      |
| 得出结果         |                |               |
| 0.001 187 76 | -0.003 226 237 | 15.177 826 98 |

由此得到近似价值函数的线性表达式如下:

$$\bar{V}_t^n = 0.0012(R_{a'}^n)_{a'} - 0.0032(M_{b'}^n)_{b'} + 15.1778 \quad (6)$$

这里采用粒度比较大的线性拟合方式,拟合前这组系统状态近似值的空间表现形式和拟合后近似价值函数的空间展现形式分别如图 4 和图 5 所示,由于线性函数存在的误差较大,因此本文用尽可能多的离散值,用非线性的表达方式来得出这个函数。

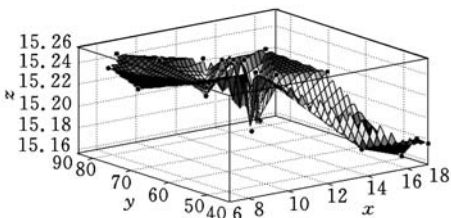


图 4 拟合前达到稳态的系统状态近似值空间分布

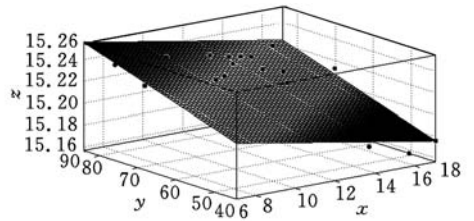


图 5 拟合后近似价值函数的空间形式

图 4、图 5 中,  $z$  轴为当期系统系统状态近似值,  $x$  轴为当期车辆动用数量,  $y$  轴为当期任务满足数量,由图可见近似价值函数的线性表达式能够比较好地匹配解空间的值。

#### 4.2 算法正确性验证

得到了近似价值函数的表达式之后,我们首先进行算法正确性的验证。利用单期决策(忽略一期以后的影响)的满意度“ $D$ ”来反映算法的正确性,决策满意度的计算如下:

$$D \approx N_1 / N_2 \quad (7)$$

其中,  $N_1$  表示当期应该被满足的运输请求任务数,  $N_2$  表示应用近似价值函数计算后得到的当期实际被满足的任务数。决策满意度越高,说明算法越正确。

通过近似价值函数的表达式(式(6))求解决策函数(式(1)),得到的调度决策结果为  $x_{1ad} = 5$ ,  $x_{2ad} = 1$ ,即 1 时刻的任务全部派遣车辆执行,2 时刻的任务执行任务 6。

算法正确性分析:最优的调度决策为 1 时刻和 2 时刻的 7 个任务应该全部执行,即  $N_1 = 7$ ,而近似价值函数求解决策函数给出的调度决策是实际执行 6 个任务,即  $N_2 = 6$ 。如果下一时刻还能满足条件的话,会延期调度剩余的任务。

决策满意度由式(7)计算为 85.71%,即正确性为 85.71%。可见,算法能够在较短时间内,得到正确性较高的近似满意解。

#### 4.3 算法优越性验证

为了进行算法的优越性验证,首先从算法的策略角度进行分析。基于 ADP 的大规模、多类型车辆动态调度算法的优越性,主要体现在算法在每期的调度决策中不但都考虑了当期决策对当期系统状态的影响,还考虑了当期决策对系统未来各期的影响。由此,如果我们仅以基于 ADP 算法中的启发式规则集为基础,只考虑当期决策对当前系统贡献值的影响而不考虑对将来各期的影响,设计一个大规模、多类型车辆动态调度的贪心算法,贪心算法实现就是任务到达只要有车辆满足条件就立即调度,这样就可以比较出两种调度策略的差异,从而判断哪种调度策略更为优越。

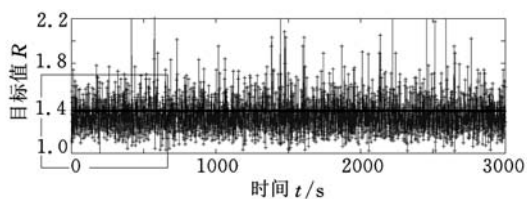
为了评判两种调度策略的优劣,我们根据问题的目标函数定义如下评判指标:

$$R = \frac{N^{(c)}}{N^{(v)}} \quad (8)$$

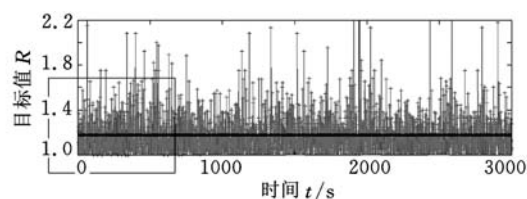
其中, $R$  为目标值, $N^{(c)}$  为每期被满足的任务数, $N^{(v)}$  为每期调度动用的车辆数。对于本文的问题,我们的优化目标是:对于每期调度决策,在尽可能完成任务的前提下,动用车辆数最少。那么,式(8)中的目标值越大,则表明每期执行相同任务数的前提下,车辆动用的数量越少。

对两种算法给定同一组算例参数:取样次数  $N$  为 100,训练周期  $T$  为 30,每期任务数为 1~15 的随机数。

经过运行后,两种算法的各期目标值的平均值的整体图和局部图分别如图 6 和图 7 所示。

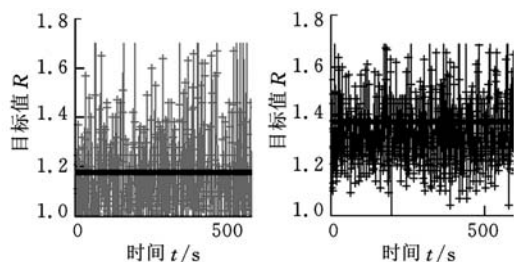


(a) 基于启发式的 ADP 算法



(b) 基于启发式的贪心算法

图 6 两种算法的目标值比较(整体图)



(a) 图 6a 放大图

(b) 图 6b 放大图

图 7 两种算法的目标比较(局部放大图)

由图 6 和图 7 可见,基于 ADP 的算法策略目标值的平均值在 1.4 左右,而贪心算法目标值的平均值在 1.2 左右,这表明,按照 ADP 算法策略进行车辆调度,任务完成数量与车辆动用数量比值的平均值要比按照贪心算法策略高 16.7%,即对于同样的任务,按 ADP 的调度策略进行车辆调度比按照贪心策略调度进行车辆调度,平均每期的车辆动用数量要少 16.7%。这说明,既考虑每期调度决策的当期贡献值,又考虑对未来各期影响的 ADP 策略,要比只考虑当期贡献值的贪

心策略优越,从而验证了算法的优越性。

## 5 结语

车辆调度问题因其规模大、类型多、不确定性强等特征需要动态的调度模型与调度优化算法。针对大规模、多类型车辆动态调度问题,基于近似动态规划的思想建立了系统状态模型和调度决策模型,提出了车辆动态调度的一系列启发式规则,在此基础上,提出了基于 ADP 的车辆动态调度训练算法和应用算法。基于训练算法所获得的仿真数据和属性聚集获得了系统的近似价值函数,基于该近似价值函数在应用阶段能够在线快速生成调度决策,为实际决策提供依据。实验结果表明,近似价值函数能够形成对状态价值的有效近似,算法(在缺少未来信息的前提下)所生成的调度能够兼顾当前和未来,显著好于贪心算法。因此,所提出模型和方法是解决大规模、多类型、不确定车辆动态调度问题的有效思路。实际应用过程中,其效果受模型的质量、参数的辨识准确性、训练数据的可获得性及准确性、对系统状态演化的实时跟踪能力等影响,需要针对真实系统进行适应性的调整、仿真和近似。

## 参考文献:

- [1] Gendreau M, Potvin J Y. Dynamic Vehicle Routing and Dispatching[J]. Fleet Management and Logistics, 1998, 115-126.
- [2] Powell W B. Stochastic and Dynamic Networks and Routing[J]. Department of Civil Engineering and Operations Research, Program in Statistics & Operations Research, 1995, 8: 141-295.
- [3] Minkoff A S. A Markov Decision Model and Decomposition Heuristic for Dynamic Vehicle Dispatching[J]. Operations Research, 1993, 41: 77-90.
- [4] 张景玲, 赵燕伟, 王海燕, 等. 多车型动态需求车辆路径问题建模及优化[J]. 计算机集成制造系统, 2010, 16(3): 544-550.
- Zhang Jingling, Zhao Yanwei, Wang Haiyan, et al. Modeling and Algorithms for a Dynamic Multi-vehicle Routing Problem with Customers' Dynamic Requests[J]. Computer Integrated Manufacturing Systems, 2010, 16(3): 544-550.
- [5] 王旭, 葛显龙, 代应. 多车型动态需求车辆路径问题建模及优化[J]. 控制与决策, 2012, 27(2): 175-181.
- Wang Xu, Ge Xianlong, Dai Ying, Research on Dynamic Vehicle Routing Problem Based on Two-phased Algorithm[J]. Control and Decision, 2012, 27(2): 175-181.

(下转第 693 页)

- [6] Fleck N A, Muller G M, Ashby M F, et al. Strain Gradient Plasticity: Theory and Experiment[J]. Acta Metall. Mater., 1994, 42: 475-487.
- [7] Stolken J S, Evans A G. A Microbend Test Method for Measuring the Plasticity Length Scale[J]. Acta Metall. Mater., 1998, 46: 5109-5115.
- [8] Aifantis E C. On the Microstructural Origin of Certain Inelastic Models[J]. Mater. Engng. Technol., 1984, 106: 326-330.
- [9] Aifantis E C. Strain Gradient Interpretation of Size Effects[J]. International Journal of Fracture, 1999, 95(1/4): 299-314.
- [10] Lam D C C, Yang F, Chong A C M, et al. Experiments and Theory in Strain Gradient Elasticity[J]. Mech. Phys. Solids, 2003, 51: 1477-1508.
- [11] 黄克智, 邱信明, 姜汉卿. 应变梯度理论的新进展(一)偶应力理论和 SG 理论[J]. 机械强度, 1999, 21(2): 81-87.  
Huang Kezhi, Qiu Xinming, Jiang Hanqing. Recent Advances in Strain Gradient Plasticity- I —— Couple Stress Theory and SG Theory[J]. Journal of Mechanical Strangth, 1999, 21(2): 81-87.
- [12] 李河宗, 董湘怀, 王倩, 等. CuZn37 黄铜板料微塑性成形中的尺寸效应研究[J]. 材料科学与工艺, 2011, 19(4): 15-19.  
Li Hezong, Dong Xianghuai, Wang Qian. Size Effects of CuZn37 Brass Foil in Microforming[J]. Materials Science and Technology, 2011, 19(4): 15-19.
- [13] 聂志峰, 周慎杰, 韩汝军, 等. 应变梯度弹性理论下微构件尺寸效应的数值研究[J]. 工程力学, 2012, 29(6): 38-46.  
Nie Zhifeng, Zhou Shenjie, Han Rujun, et al. Numerical Study on Effects of the Microstructures Based on Strain Gradient Elasticity[J]. Engineering Mechanics, 2012, 29(6): 38-46.
- [14] 周丽, 李守新, 黄树涛. 运用应变梯度塑性理论模拟颗粒增强铝合金强度及延伸率的尺寸效应[J]. 功能材料, 2008, 39(12): 2101-2105.  
Zhou Li, Li Shouxin, Huang Shutao. Simulation of Size Effect on Strength and Elongation of Al Alloy Wity Rarticles Reinforcement Using Strain Gradient Plasticity Concept [J]. Journal of Functional Materials, 2008, 12(39): 2101-2105.
- [15] Fleck N A, Hutchinson J W. A Phenomenological Theory for Strain Gradient Affects in Plasticity [J]. Journal of the Mechanics and Physics of Solids, 1993, 41: 1825-1857.
- [16] Fleck N A, Hutchinson J W. Strain Gradient Plasticity[J]. Advanced Applied Mechanics, 1997, 33: 295-361.
- [17] Saha R, Xue Z, Huang Y, et al. Indentation of a Soft Metal Film on a Hard Substrate: Straingradient Hardening Effects[J]. Mech. Phys. Solids, 2001, 49: 1997-2014.
- [18] Liu Z L, Zhuang Z, Liu X M, et al. A Dislocation-dynamics Based Higher-order Crystal-plasticity Model and Applications on Confined Thin-film Plasticity[J]. Int. J. Plast., 2011, 27: 201-216. (编辑 王艳丽)
- 
- 作者简介: 郭幼丹, 男, 1960 年生. 集美大学机械与能源工程学院副教授. 主要研究方向为精密成形技术. 发表论文 50 余篇.  
程晓农, 男, 1958 年生. 江苏大学材料科学与工程学院教授、博士研究生导师.
- ~~~~~
- (上接第 688 页)
- [6] 袁建清. 基于 TS 的动态车辆调度问题的混合算法研究[J]. 计算机现代化, 2012(6): 73-75.  
Yuan Jianqing. Mixed Tabu Search Algorithm for Dynamic Vehivle Scheduling Problem [J]. Jisuanji Yu Xiandaihua, 2012(6): 73-75.
- [7] Azi N, Gendreau M, Potvin J Y. A Dynamic Vehicle Routing Problem with Multiple Delivery Routes [J]. Annals of Operations Research, 2012, 199(1): 103-112.
- [8] Warren B. Powell, Approximate Dynamic Programming for Operations Research[M]. Princeton: Department of Operations Research and Financial Engineering Princeton University, 2005.
- [9] Clara Novoa, Robert Storer. An Approximate Dynamic Programming Approach for the Vehicle Routing Problem with Stochastic Demands[J]. European Journal of Operational Research, 2009, 196(2): 509-515.
- [10] Si J, Barot A G, Powell W B, et al. Handbook of Learning and Approximate Dynamic Programming: Sealing up to the Real World [M]. New York: IEEE Press and John Wiley & Sons, 2004.
- [11] Kirk D E. Optimal Control Theory: An Introduction[M]. Englewood Cliffs: Prentice\_Hall, 1970. (编辑 王艳丽)
- 
- 作者简介: 李 雷, 女, 1989 年生. 哈尔滨工业大学计算机科学与技术学院硕士研究生. 主要研究方向为信息物理系统、资源动态调度.  
聂兰顺, 男, 1979 年生. 哈尔滨工业大学计算机科学与技术学院副教授.  
齐文艳, 女, 1987 年生. 哈尔滨工业大学计算机科学与技术学院硕士研究生.  
战德臣, 男, 1965 年生. 哈尔滨工业大学计算机科学与技术学院教授、博士研究生导师.