

基于滚动时域优化策略和深度 Q 学习算法的 AGV 充电调度方法

王志杰

杭叉集团股份有限公司 浙江杭州

【摘要】设计基于滚动时域优化策略和深度 Q 学习算法的 AGV 充电调度方法。首先，在滚动时域优化策略下，确定了 AGV 的充电状态，控制 AGV 的充电负荷。其次，建立 AGV 系统当前状态的深度 Q 学习算法模型，获取 AGV 执行相应动作的 Q 值。最后，计算 AGV 充电调度功耗，实现 AGV 的高效充电调度。

【关键词】Q 学习算法；AGV；调度方法

【收稿日期】2023 年 1 月 25 日 **【出刊日期】**2023 年 2 月 20 日 **【DOI】**10.12208/j.ijme.20230006

AGV charging scheduling method based on rolling time domain optimization strategy and deep Q learning algorithm

Zhijie Wang

HANGCHA GROUP CO., LTD. Hangzhou, Zhejiang

【Abstract】An AGV charging scheduling method based on rolling time domain optimization strategy and deep Q learning algorithm was designed. Firstly, under the rolling time domain optimization strategy, the charging state of AGV is determined and the charging load of AGV is controlled. Secondly, the deep Q learning algorithm model of the current state of the AGV system is established to obtain the Q value of the AGV executing corresponding actions. Finally, the power consumption of AGV charging scheduling is calculated to achieve efficient charging scheduling of AGV. Key words: rolling time domain optimization strategy; Deep Q learning algorithm; AGV. Charge scheduling method

【Keywords】Q learning algorithm; AGV. Scheduling method

引言

基于多目标优化模型的调度方法中，建立的多目标动态优化模型，控制了充电功率与充电时间，消减了一部分负荷峰谷值的影响，有效地节约了充电成本。基于 Stackel berg 博弈的调度方法中，建立了一个 Stackel berg 博弈调度模型，使充电负荷存在唯一一个 Stackel berg 均衡点，有效地约束了充电负荷，使充电效益最大化。但是，以上两种充电调度方法中，均以充电调度模型为主，仅能够保证 AGV 的有序充电，充电功率始终较高，与 AGV 的经济需求不符。滚动时域优化是一种特定的控制方法，以滚动式优化的方式，为相关物质提供时刻优化的性能指标。深度 Q 学习算法是一种数值迭代的调度规划方法，能够为相关领域提供良好的优化动作。因此，本文利用滚动时域优化策略与深度 Q 学习算法，

设计了 AGV 充电调度方法。

1 滚动时域下基于 Q 学习的 AGV 充电调度方法设计

1.1 基于滚动时域优化策略确定 AGV 的充电状态

AGV 系统的充电状态包括充电、搬运、休息等状态，本文将 AGV 的充电调度问题看做一个多目标优化的问题，将 AGV 的电池电量约束、充电功率、充电时间、充电负荷均等目标进行综合优化，提高 AGV 的运输水平^[1]。本文设计的 AGV 充电调度总流程如下图 1 所示。

如图 1 所示，本文按照上述三个步骤进行充电调度。其中，Q 值是学习算法模型的期望值，本文将 AGV 的最大期望值设定为总行驶时间。在 AGV 充电过程中，充电状态的确定通过状态函数评定，

作者简介：王志杰（1984-）男，浙江萧山，汉，工程师，无人驾驶工业车辆研发及设计

将充电状态与报酬期望相对应，即可得出当前 AGV 状态的函数，函数表达式如下式 (1)：

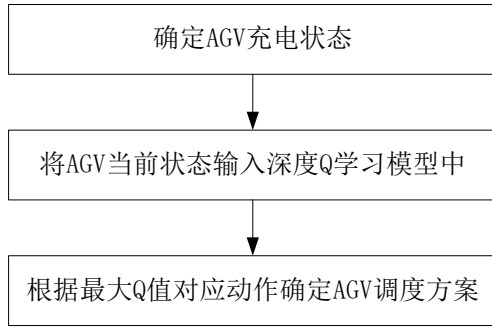


图 1 基于深度 Q 学习的调度流程图

$$V^*(s) = \max_{\pi} E \left[\sum_{t=0}^H \lambda^t R(s_t, A_t, s_{t+1}) \mid \pi, s_0 = s \right] \quad (1)$$

式 (1) 中， $V^*(s)$ 为 AGV 的当前充电状态函数值； E 为 AGV 最终状态的回报期望； λ^t 为折扣因子，其在 0~1 的范围内变化， λ^t 值越小，充电状态越正确； H 为充电完成时刻； R 为 AGV 下一状态执行的搬运任务； t 为某一特定的时刻； s_t 为 t 时刻下的充电状态； A_t 为 t 时刻的对应权重； s_{t+1} 为 $t+1$ 是时刻下的充电状态； s_0 为 AGV 初始电量； s 为初始期望。为了使 AGV 的充电状态更加明确，本文在滚动时域优化条件下，对 AGV 系统进行充电负荷控制^[2]。在 t 时刻，根据已知状态函数 $V^*(s)$ ，得出滚动时域极小化代价函数为：

$$K(t) = \left\| \hat{x}(k-E) - \bar{x}(k-E) \right\|^2 \quad (2)$$

式 (2) 中， $K(t)$ 为滚动时域极小化代价函数； \hat{x} 为充电负荷峰谷差值； \bar{x} 为充电负荷平均值。当充电负荷较大时，降低 E ；当充电负荷较小时，提高 E ，以此完成对 AGV 系统的充电负荷控制。

1.2 建立 AGV 系统当前状态的深度 Q 学习算法模型

在 AGV 形式的某一时刻，在状态 s 下进行 S 动作，获得报酬的过程，是深度 Q 学习算法的原理。本文根据 AGV 当前充电状态，建立 AGV 调度框架模型，如下图 2 所示。

如图 2 所示，本文将 S20~S28 作为深度 Q 学习算法模型的步骤。在模型接收开始搬运的指令之后，将 AGV 系统状态进行初始化处理，将 AGV 系统中

的相关参数恢复出厂设置。此时，将模型与 AGV 建立一个状态 s ，根据 s 选择 AGV 的搬运动作，若充电电量能够满足其完成该动作，则任务结束；若充电电量无法满足其完成该动作，则更新判定 Q 值，再次选择一个搬运动作进行 AGV 作业，重复上一步骤，直至任务结束。本文在深度 Q 学习过程中，将 AGV 作业最大期望值作为 Q 值，则 Q 值的最优函数表达式如式 (3) 下：

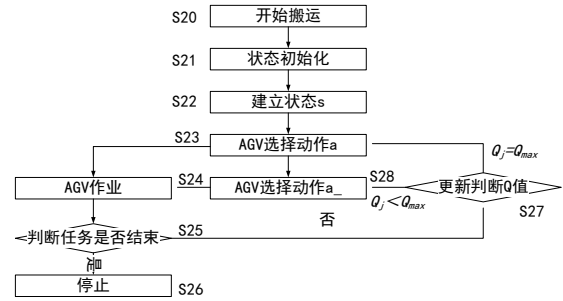


图 2 AGV 调度框架模型图

$$Q^*(s, a) = \sum_{s'} P(s'|a) (R(s, a, s') + \lambda' \max_{a'} Q^*(s', a')) \quad (3)$$

式 (3) 中， $Q^*(s, a)$ 为 Q 值的最优函数表达式； s' 为 AGV 的下一次搬运动作； $P(s'|a)$ 为下一次搬运动作获得的最大回报； $R(s, a, s')$ 为任务结束获得的最大回报； a' 为模型与 AGV 建立的下一状态； $Q^*(s', a')$ 为下一状态下的 Q 值最优函数。在此状态下，本文对 Q 值进行更新，公式如下：

$$V(Q) = V^*(s) + \beta (R_{t+1} + \lambda' Q^*(s, a) - V^*(s)) \quad (4)$$

式 (4) 中， $V(Q)$ 为更新后的 Q 值； β 为深度学习因子， $0 < \beta < 1$ 。当更新后的 Q 值 > 更新前的 Q 值时，表示该 Q 值能够对应相应的 AGV 动作；当更新后的 Q 值 < 更新前的 Q 值时，表示该 Q 值无法对应相应工作，需要再次更新。

1.3 计算 AGV 充电调度功耗

为了实现 AGV 充电的高效调度，本文在确定 AGV 充电状态、建立算法模型的基础上，对 AGV 的充电调度功耗进行计算，在确定 AGV 的充电状态的基础上，进一步控制充电负荷。每一个 AGV 中均存在一个较大的电能储备量，能够保证 AGV 的正常行驶。每次 AGV 到达充电站时，其充电时间、充电功率、充电负荷等指标均需要进行调度控制，以此保证 AGV 的正常使用。在 AGV 充电时间段内，其

调度目标函数为:

$$f(x) = \min \sum_{i=1}^H P_{i,t} \quad (5)$$

式(5)中, $f(x)$ 为 AGV 的调度目标函数; $P_{i,t}$ 为负荷最小峰值。将调度目标节点的充电负荷作为标准指标, 则 AGV 调度一次的充电功耗为:

$$P_c = \min \left[\sum_{i=1}^H \sum_{t=1}^E (U_{i,t} / I_{i,t} \cdot h) \right] \quad (6)$$

式(6)中, P_c 为 AGV 调度一次的充电功耗; $U_{i,t}$ 为 AGV 调度一次的最小电压; $I_{i,t}$ 为 AGV 调度一次的最小电流; h 为充电时间。通过计算充电调度功耗, 能够更加直观地分析该调度方法的经济效果, P_c 越低, 该调度方法的经济效益越高, 越能够符合 AGV 的充电调度需求。

2 实验

为了验证本文设计的充电调度方法是否具有使用价值, 本文对上述方法进行了实现分析。最终的实验结果以文献[1]基于多目标优化模型的调度方法、文献[2]基于 Stackel berg 博弈的调度方法, 以及本文设计的基于滚动时域与 Q 学习算法的调度方法进行对比的形式呈现。具体实验过程及实验结果如下所示。

2.1 实验过程

本次实验将 AGV 充电负荷接入中低压配电网中, 作为一种移动的负荷, 在各个供电区域为 AGV 分配合理的电能。本次实验采用按需分配的方式, 在 AGV 行驶区域设置了多个换电站, 换电站的充电功率为目标充电功率, 保证充电负荷与电网总负荷处于均匀分布的状态。充电设施的分布示意图如下图 3 所示。

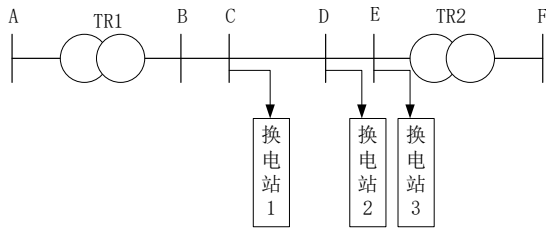


图 3 充电设施的分布示意图

如图 3 所示, A-F 的区间为 AGV 的行驶区间, C、D、E 分比三个换电站, B 为其他电动汽车使

用的负荷, TR1、TR2 是 AGV 在路上行驶时的充电节点。由此看来, AGV 在整个行驶过程中, 并未到达换电站进行充电, 而是在 A-B、D-F 之间的距离进行了充电, 需要对其进行充电调度, 为 AGV 的后续行驶提供电能。为了更好进行充电调度, 本文对深度 Q 学习算法模型的收敛性能进行考察。收敛性能作为模型的重要指标, 本文根据特征关联步长, 对模型的近似最优平均代价进行计算, 公式如下:

$$\gamma_i = \frac{0.3}{1 + \frac{\varepsilon_i}{10}} \quad (7)$$

式(7)中, γ_i 为模型的近似最优平均代价; ε_i 为特征关联步长。将深度 Q 学习算法模型的 γ_i 计算出来之后, 对其收敛性能进行分析, 如下图 4 所示。

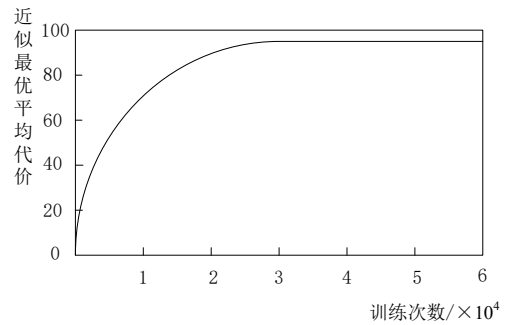


图 4 深度 Q 学习算法模型的收敛性能

如图 4 所示, 在模型训练次数为 3×10^4 时, 近似最优平均值 γ_i 最高且稳定, 此时模型已经收敛, γ_i 值不再变化。为了保证本次实验的准确性, 本文选取了另一组特征关联步长 V_i , $V_i > \varepsilon_i$, 同样对其进行了训练, 在训练次数达到 3×10^4 时, 近似最优平均值 γ_i 最高且稳定, 能够保证模型的收敛效果。本文在模型训练 3×10^4 之后, 对 AGV 进行充电调度, 分析在调度过程中, AGV 的充电功率变化情况。

2.2 实验结果

在上述实验条件下, 本文随机选取了 10%~80% 的 AGV 电池初始电量, 作为实验基础指标。并在充电时间一致的前提下, 对文献[1]基于多目标优化模型调度方法的 AGV 充电功率、文献[2]基于 Stackel berg 博弈调度方法的 AGV 充电功率, 以及本文设计的基于滚动时域与 Q 学习算法调度方法的 AGV 充电功率进行对比, 验证本文设计方法的充电调度效果。实验结果如下表 1 所示。

如表 1 所示, 本文以 10%~80%的 AGV 电池初始电量作为实验指标, 在 3h~24h 的充电时间内, 进行了充电功率分析。已知, 在同样的充电时间下, 充电功率越低, 充电调度成本越低, 经济优势越高。在其他条件均一致的情况下, 文献[1]基于多目标优化模型调度方法的 AGV 充电功率相对较高, 在 205.47kW~345.28kW 的范围内变化, 在 AGV 电池

电量为 10%、20%时, 充电功率高于最大功率, 充电调度成本随之增加, 无法满足 AGV 低功耗的充电调度需求, 导致 AGV 不能得到及时的电能补给, 影响其继续使用性能^[3]。

文献[2]基于 Stackel berg 博弈调度方法的 AGV 充电功率有所下降, 在 144.12kW~328.43kW 的范围内变化。

表 1 实验结果

AGV 电池 初始电量/%	最大充电 功率/kW	文献[1]基于多目标优化模 型调度方法的 AGV 充电功 率/kW	文献[2]基于 Stackel berg 博 弈调度方法的 AGV 充电功 率/kW	本文设计的基于滚动时域与 Q 学习算法调度方法的 AGV 充电 功率/kW
10	326.52	345.28	328.43	158.63
20	308.47	312.63	256.42	132.74
30	298.96	286.72	243.55	112.12
40	278.63	275.43	232.12	108.45
50	257.56	256.58	222.47	95.43
60	246.45	232.14	168.12	88.67
70	222.36	211.62	152.18	69.55
80	208.16	205.47	144.12	54.13

在 AGV 电池电量为 10%时, 充电功率稍高于最大功率, 充电调度成本同样较高。而本文设计的基于滚动时域与 Q 学习算法调度方法的 AGV 充电功率相对较低, 在 54.13kW~158.63kW 的范围内变化, 均在最大充电功率以内, 能够满足 AGV 低功耗的充电调度需求, 及时补给 AGV 电能, 为 AGV 的稳定运行提供保障, 符合本文研究目的。

3 结束语

近些年来, AGV 的使用较为广泛, 为物流运输提供了更加灵活的工作环境。由于仓储装卸的作业量较大, 通常需要多个 AGV 共同工作, 保证装卸工作的圆满完成, 由此造成的充电调度问题亟待解决。因此, 本文在滚动时域优化策略下, 设计了基于深度 Q 学习的 AGV 充电调度方法。通过深度 Q 学习模型、充电状态、调度功耗等方面, 降低了 AGV 调度的充电功率, 为 AGV 提供低成本的使用环境, 提高 AGV 的工作效率^[4-5]。

参考文献

- [1] 刘岩,邓彬,王瑾,等.基于多目标优化模型的电动汽车充电调度策略[J].沈阳工业大学学报,2022,44(02):127-132.
- [2] 冯小,张传林,崔承刚,等.基于 Stackelberg 博弈的孤岛式光储充电站调度优化[J].电网技术,2022,46(10):3989-4001.
- [3] 张彬桥,李成,李振兴,等.基于交通信息和配电网全成本电价的电动汽车充电负荷调度方法[J].智慧电力,2021,49(07):52-57.
- [4] 季娜,蔡红霞,钱晖.基于滚动时域优化策略的玻璃深加工车间调度研究[J].工业控制计算机,2017,30(3):3.
- [5] 丁一陈婷.基于滚动时域优化策略的多载 AGV 充电调度[J].中国航海,2020,43(2):80-85.

版权声明: ©2023 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS