

基于蒙特卡洛树搜索算法实现轨道交通车辆多功能 车辆总线周期调度表优化

耿力¹ 耿强²

(1. 同济大学铁道与城市轨道交通研究院, 201804, 上海; 2. 电子科技大学自动化工程学院, 611731, 成都//第一作者, 硕士研究生)

摘要 目的: 为改善轨道车辆 MVB(多功能车辆总线)周期调度表的负载均衡度, 提高 MVB 带宽利用率和车辆的通信系统实时性, 需要对 MVB 周期调度表进行优化。方法: 简述了 MVB 周期信息通信原理, 明确了 MVB 调度表的 2 个约束条件, 建立将 MCTS(蒙特卡洛树搜索)算法应用于 MVB 调度表优化的理论方法, 并通过 Python 软件编程实现该算法。明确了 MCTS 算法的优化目标, 结合 MVB 调度表的约束条件设计了优化 MCTS 算法的预剪枝策略。选取了 RMS(单调速率调度)算法、MCTS 算法和 GA(遗传算法)三种算法进行仿真试验, 并对各算法的结果进行对比分析。结果及结论: 优化后的 MCTS 算法有效避免了大量无效的搜索, 搜索效率非常高。与 RMS 算法、GA 相比, 优化后的 MCTS 算法能在相同的搜索时间内获得更均衡的解。如果车辆通信设备增加, 需要生成包含更多变量的调度表时, MCTS 算法更能凸显其搜索优势。

关键词 轨道车辆; 多功能车辆总线; 周期调度表; 蒙特卡洛树搜索算法; 强化学习

中图分类号 U270.38⁺2

DOI: 10.16037/j.1007-869x.2023.11.013

are clarified, and a pruning strategy for optimizing the MCTS algorithm is designed based on the constraints of the MVB scheduling table. Three algorithms, RMS (Monotonic Rate Scheduling) algorithm, MCTS algorithm, and GA (Genetic Algorithm), are selected for simulation experiments, and the results of each algorithm are compared and analyzed. **Result & Conclusion:** The optimized MCTS algorithm effectively avoids a large number of invalid searches with a very high search efficiency. Compared with RMS algorithm and GA, the optimized MCTS algorithm can obtain more balanced solutions within the same search time. If the number of vehicle communication devices increases and a scheduling table containing more variables needs to be generated, the advantages of the MCTS algorithm in searching can be further highlighted.

Key words rail transit vehicle; multi-function vehicle bus; periodically scheduling table; Monte Carlo tree search approach; reinforcement learning

First-author's address Institute of Rail Transit, Tongji University, 201804, Shanghai, China

Optimization of Multi-Functional Vehicle Bus Scheduling Table Based on Monte Carlo Tree Search Algorithm

GENG Li, GENG Qiang

Abstract **Objective:** In order to improve the load balance of the MVB (multi-functional vehicle bus) periodic scheduling table for rail vehicles, increase the utilization of MVB bandwidth, and achieve the real-time performance of the vehicle's communication system, it is necessary to optimize the MVB periodic scheduling table. **Method:** The principle of MVB periodic information communication is briefly introduced, the two constraints of MVB scheduling table are clarified, a theoretical method for applying MCTS (Monte Carlo Tree Search) algorithm to MVB scheduling table optimization is established, and the algorithm through Python software programming is implemented. The optimization objectives of the MCTS algorithm

轨道车辆通信总线由 WTB(绞线式列车总线)和 MVB(多功能车辆总线)构成。其中: WTB 主要用于日常作业时经常改变编组数量的列车间连接各车辆的串行数据总线; MVB 主要用于固定车厢或具有固定编组列车这一特定范围的轨道车辆的通信网络^[1]。MVB 将一个车厢内或一个编组内的可编程设备互连, 并可直接连接简单的传感器和执行机构, 从而为车厢内各设备的诸多功能(如车门控制、列车制动、空调调节、发布旅客信息及预留坐席等)的自动实现、消息的传达、资源的共享, 以及各设备间的合理调配提供可靠且顺畅的信息交换通道。在车辆上有众多设备, 设备之间对 MVB 的占用通过 MVB 周期调度表(以下简称“调度表”)来控制。合理的调度表能提高 MVB 的带宽利用率, 并通过缩短通信周期来提高车辆通信系统的实

时性。

调度表的设计主要有两种方案。方案一用经典的 RMS(单调速率调度)算法设计调度表, RMS 算法的优点是速度快, 缺点是其调度表不具有负载均衡特性, 易引起带宽浪费, 其通信的实时性也较低。方案二借助启发式算法找出合理的调度顺序。因为寻找负载均衡的调度方式是一个 NP-Hard(非确定性多项式困难)的组合优化问题, 其组合方式随变量的数目呈指数增长, 例如针对 20 个变量的调度表设计, 普通的计算机已经难以穷举求解, 通常采用启发式算法求解。已提出用于调度表设计的启发式算法包括 GA(遗传算法)^[2]、混合遗传算法^[3]、免疫遗传算法^[4]、蚁群算法^[5]及改进的差分进化算法^[6]等, 这些算法的相似度高, 很多算法属于对同一算法进行了不同程度的改进, 其缺点是搜索过程中可能产生大量效果很差的解, 进而造成搜索效率低。部分改进过后的算法将搜索分成两个阶段进行^[6], 第一阶段为全局搜索, 第二阶段为局部搜索, 这在一定程度上提高了搜索的效率, 但也面临如更多超参数等问题。

基于强化学习的围棋人工智能程序 Alpha-Go 采用 MCTS(蒙特卡洛树搜索)算法进行启发式搜索^[7], 该算法很好地解决了组合优化问题中搜索空间巨大、难以搜索到优质解这一难题。因此, 本文尝试将 MCTS 算法引入 MVB 调度表设计中, 并针对 MVB 调度表的特点对该算法进行改进, 以提高搜索效率。

1 MVB 周期信息通信原理

1.1 MVB 调度表及约束条件

MVB 上有多个节点, 主节点根据调度表的规则将特定的从设备指定为源节点, 赋予该源节点发送数据的能力。该调度过程不断循环, 将每个循环周期称为宏周期。为了能够传输非周期信息, 主节点将时间等分成若干个时间片, 并将该时间片称为微周期。每个微周期被划分成两部分, 第一部分称为周期相, 用于传输周期性数据; 第二部分称为偶发相, 用于传输非周期性数据。国际电工委员会发布的 IEC 61375 标准第 3-1 部分对 MVB 调度表的宏周期及微周期中的周期相时间占比等条件进行了约束。使用各种算法生成的调度表, 需要检查是否满足规定的约束条件。若满足约束条件, 则称该调度表是可调度的, 否则称该调度表是不可调度的。

MVB 的约束条件包括约束条件 1 及约束条件 2。

1.1.1 约束条件 1

约束条件 1 为变量首次出现位置必须不超过其周期, 可形式化表示为

$$\forall i \in \{1, 2, \dots, m\}, p_i \leq T_i / T_{bp} \quad (1)$$

式中:

i ——变量编号;

m ——变量总数;

T_{bp} ——微周期, 单位 ms;

p_i ——变量 i 首次出现的周期位置;

T_i ——变量 i 的周期, 单位 ms。

1.1.2 约束条件 2

约束条件 2 为根据 IEC 61375 标准第 3-1 部分, 每个微周期中周期相所占时间最多为微周期的 65%, 即

$$\forall j \in \{1, 2, \dots, z\}, U_j \leq 0.65 T_{bp} \quad (2)$$

式中:

j ——微周期的编号;

z ——微周期总数;

U_j ——第 j 个微周期中周期相所占时间, 单位 ms。

1.2 优化目标

算法优化的目标为: 在一个宏周期中各微周期内的周期相时间尽可能均衡。记一个组合方式为 G , 该组合方式对应一个调度表。设 \bar{U} 为微周期中周期相所占时间的平均值, 使用标准差 $\sigma(G)$ 来衡量 MVB 负载的均衡度。 $\sigma(G)$ 越小, 则 MVB 负载越均衡。 $\sigma(G)$ 的数学表达式为:

$$\sigma(G) = \sqrt{\frac{1}{z} \sum_{j=1}^z (U_j - \bar{U})^2} \quad (3)$$

$$\bar{U} = \frac{1}{z} \sum_{j=1}^z U_j \quad (4)$$

算法的优化目标转为找到最佳的组合方式 G^* , 使得 $\sigma(G)$ 最小, 即:

$$G^* = \arg \min_G \sigma(G) \quad (5)$$

2 MCTS 算法原理简述

本文首次将强化学习中的 MCTS 算法引入到 MVB 调度表的构建过程中。MCTS 算法是一种在决策空间进行随机采样并构建出一棵搜索树, 通过搜索树来实现寻找最优决策的算法。MCTS 算法在序列决策问题、博弈问题及规划问题中均具有重要

的影响力,特别是在围棋 AI(人工智能)程序中的成功应用,使得该算法备受关注。MCTS 算法原理如图 1 所示。

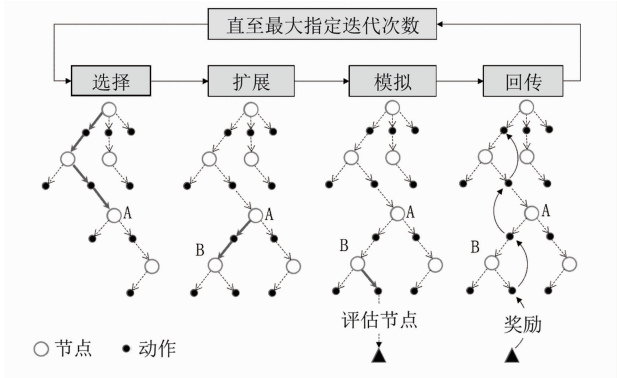


图 1 MCTS 算法原理示意图

Fig. 1 Schematic diagram of MCTS approach

如图 1 所示,MCTS 是一个不断迭代的过程,每次迭代可分为 4 个步骤:

1) 步骤 1:选择。从根节点开始,按照一定的策略递归向下伸展,直到抵达最需要扩展的节点处。如果一个节点没有到达终止状态,并且还有未访问的子节点,则该节点是可扩展的。如图 1 中的 A 节点。

2) 步骤 2:扩展。根据每个状态可选择的动作,对选择出的叶节点新增一个或多个子节点,从而实现树的扩展。如图 1 中的 A 节点,可将其扩展得到 B 节点。

3) 步骤 3:模拟。将步骤 2 扩展得到的新节点按照预定的方式进行演算,进而产生一个输出结果,直至游戏结束。如对 B 节点进行快速模拟时,其每个动作通常为随机选择。在游戏结束时,根据游戏规则可获得对应的奖励。

4) 步骤 4:回传。将步骤 3 模拟后获得的奖励回传,更新执行模拟节点及它的所有祖先节点的统计值。

在上述迭代过程中,步骤 1 中如何选择合适的子节点是一个关键问题,这也正是强化学习必须要面对的探索-利用困境。MCTS 算法中使用了 UCT (针对树的置信上界)方法,以平衡探索和利用。通常选择在 UCT 值最大的节点上进行扩展,其表达式为:

$$Q(v) = \frac{R(v)}{N(v)} + 2C_p \sqrt{\frac{2\ln N(p_v)}{N(v)}} \tag{6}$$

式中:

- $Q(v)$ ——节点 v 的 UCT 值;
- $R(v)$ ——节点 v 获得的累计奖励;
- $N(v)$ ——节点 v 的总访问次数;
- p_v —— v 的父节点;
- $N(p_v)$ —— p_v 被访问的总次数;
- C_p ——平衡探索与利用的常数,本文取 0.1。

式(6)等式的右边为 2 项之和,其中:第 1 项表征了节点 v 平均获得的奖励;第 2 项中,在父节点 p_v 多次被访问而某个子节点长期未被访问时,其对应的 $N(v)$ 较小,此时 $Q(v)$ 较大,则节点 v 应被访问。

3 MCTS 算法在 MVB 调度表中的优化

上文介绍了 MCTS 算法的通用框架,将其应用于 MVB 调度表时需要进行优化设计。

3.1 状态空间与动作空间

MVB 调度表的状态为当前变量的排布情况,其状态采用各个变量首次出现的微周期序号形成的整数数组来表示,所有状态的可能集合构成状态空间。不同的状态可以执行不同的动作,即选出某个未安排的变量,将其排布在某一个微周期内,其所有可执行的动作将构成动作空间。所有变量均排布完毕后,动作空间变为空集,算法结束。

3.2 奖励规则设计

在强化学习中,智能体的目标是获得尽可能多的奖励。根据 MVB 变量调度的过程,提出将当前调度表的均衡程度与历史数据进行对比,根据“进步”程度给出奖励。其奖励规则如表 1 所示。

表 1 MVB 调度表中 MTCS 算法的奖励规则

Tab. 1 The reward rules of MCTS approach for MVB scheduling table

规则序号	规则内容
1	如果当前状态高于历史最优状态,得 1 分
2	如果当前状态没有超过历史最优状态,但高于历史平均水平,得 0.1 分
3	不满足规则 1 或规则 2,得 0 分

3.3 对搜索树进行预剪枝优化

在变量排布过程中,应遵循上文所述的 MVB 调度表约束条件。通常情况下,约束条件 2 很容易满足,因此在算法实现中应先考虑约束条件 1,待得到一个较好的解时,再检查是否满足约束条件 2。约束条件 1 能过滤掉大量不满足要求的劣质解。例如,仿真时选取 20 个变量及 32 个微周期工况,在无任何约束时调度表的组合数为 20^{32} (即 4.3×10^{41}),

而满足约束条件 1 的组合数约为 5.8×10^{17} , 即只有大约 $1/10^{24}$ 的组合能够满足约束条件 1。

因此, 针对 MCTS 算法和约束条件 1, 本文设计的预剪枝规则为: 如果某个状态对应的调度表不可调度, 则将其标记为不合法节点, 以后不再对该不合法节点进行任何的扩展或模拟, 以避免大量不必要的搜索。

3.4 设置变量优先级

变量优先级的设计对提高 MVB 的均衡度有一定的影响。如果最后安排的变量传输时间很长, 则该变量的排布很可能导致整个调度表不均衡, 试验中也验证了此结论的正确性。因此, 优先级规则设定为: 变量传输时间越长, 此变量的优先级越高。

4 仿真试验及结果分析

选取 RMS 算法、MCTS 算法和 GA 三种算法进行仿真试验, 并对各算法的结果进行对比分析。其中 GA 的形式随具体应用场景的改变而略有差异, 因此先简要介绍一下本文使用的 GA 形式, 再进行算法对比。

4.1 用于 MVB 调度表的 GA

GA 是一种经典的启发式搜索算法, 其流程如图 2 所示。该算法对个体的编码方法和上文所述 MCTS 算法的编码方法一致, 即一个编码中包含了各个变量首次出现的位置, 编码长度等于变量总数 m 。初始化种群中个体数量为记作 N_p , 每轮更新中将每个个体变异。变异方式为取随机编码中的一个变量, 将其随机放到一个微周期内。如果变异后的新个体适应度大于其父亲的适应度, 则新个体代替父亲, 否则保留父亲不变, 丢弃新个体。

对 GA 而言, 适应度越大说明其对应的 MVB 负载越均衡, 因此采用式 (3) 标准差的相反数 (即 $-\sigma(G)$) 作为该算法的适应度。将种群中每个个体都变异一次叫做一次种群变异, 试验中种群变异次数记为 N_{GA} 。

4.2 算法对比分析

定义 t_i 为第 i 个变量在 1 个微周期中占用 MVB 时间。使用文献 [6] 中使用的算例数据, 其调度表包含 20 个变量及 32 个微周期, 各变量对应的周期数据如表 2 所示。

本文使用 Python 编程实现 RMS 算法、GA 和 MCTS 算法, 硬件采用普通的笔记本电脑, 其参数分

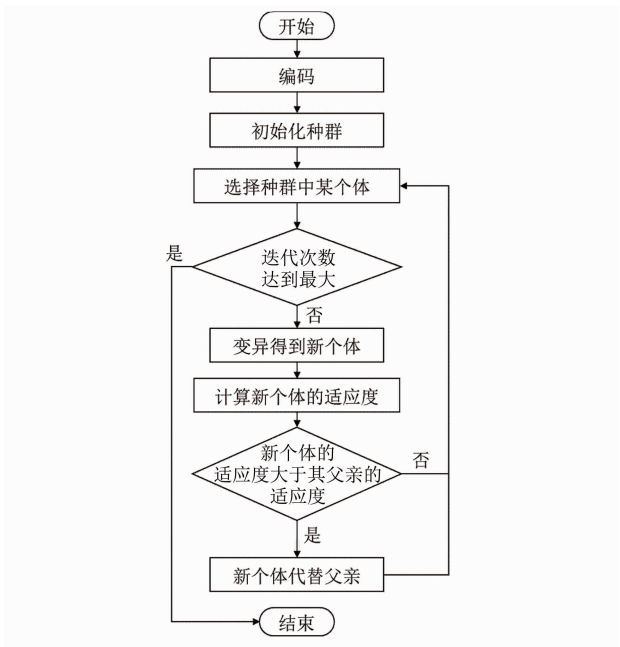


图 2 GA 流程图
Fig. 2 Flowchart of genetic algorithm

表 2 试验用的变量及其周期

Tab. 2 The variables and corresponding periods used in experiment

<i>i</i>	<i>t_i</i> /μs	<i>T_i</i> /ms	<i>i</i>	<i>t_i</i> /μs	<i>T_i</i> /ms
1	75	1	11	75	8
2	96	2	12	96	8
3	64	4	13	64	16
4	75	4	14	75	16
5	75	4	15	96	16
6	144	4	16	96	16
7	144	4	17	96	16
8	144	4	18	240	16
9	64	8	19	75	32
10	64	8	20	96	32

别为: CPU(中央处理器)的型号为 i5-8265U, 1.60 GHz 主频, RAM(随机存取存储器)的容量为 12 GiB。为便于比较 GA 和 MCTS 算法, 试验中将这两种算法的一次试验运行时间均设定为 180 min (即 10 800 s)。GA 一次试验的 $N_p = 100$ 个, $N_{GA} = 5\,000$ 万次。MCTS 算法一次试验的最大采样次数 $N_{MCTS} = 1\,500$ 万次。RMS 算法具有确定性, 但 GA 和 MCTS 算法具有一定的随机性。为了避免单次试验引起的误差, 将三种算法均重复运行 10 次。试验中发现 GA 和 MCTS 算法难以收敛, 随着时间的推移, 这两种算法可能以一定概率获得一个更优的解。因此, 不通过最终的收敛而通过一段较长时间内的搜索结果来评估 GA 和 MCTS 算法的好坏。

表 3 给出了三种算法 10 次试验后的试验平均运行时间、平均均衡度及最优解。最优解用一个向量表

达,其元素数等于 m ,向量的第 i 个值(记作 p_i)表示第 i 个变量首次出现安排在第 p_i 个微周期。

表 3 三种算法的最优搜索结果对比
Tab.3 Comparison of the optimal experimental results of three algorithms

算法	试验平均运行时间/s	平均均衡度/ μ s	算法的最优解
RMS 算法	0.002	208.3	[1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 4 4 4]
GA	10 800.000	36.3	[1 2 2 1 3 3 4 1 1 8 5 2 4 12 14 11 15 6 7 28]
MCTS 算法	10 800.000	27.2	[1 2 1 3 1 3 2 1 2 3 8 6 4 4 16 15 8 12 7 23]

试验验证可知:表 3 中各个算法给出的最优解均满足约束条件 1 和约束条件 2,即获得的调度方式是可调度的。由表 3 中还可以看出:RMS 算法下一次试验的平均运行时间最小,但均衡度最差。在长时间搜索下,MCTS 算法的搜索效率较 GA 高,其均衡度在三个算法中最小。

图 3 直观展示了 RMS 算法、GA 和 MCTS 算法三种算法生成的最优调度表每个微周期内周期相占用时间分布情况。由图 3 可知:RMS 算法生成的调度表负载严重不均衡;GA 算法生成的调度表均衡度较好;MCTS 生成的调度表负载最均衡。

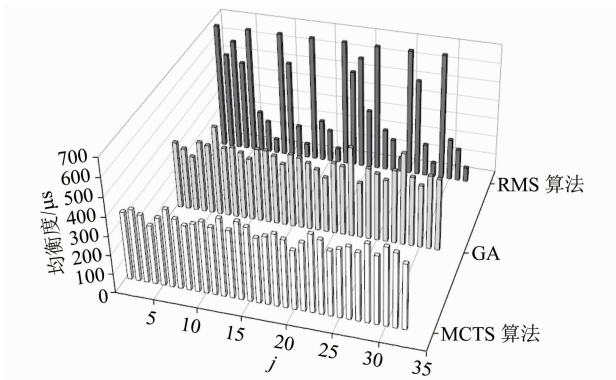


图 3 三种算法生成的最优调度表周期相占用时间分布
Fig. 3 Time distribution of periodic phases generated by three algorithms respectively in the optimal scheduling table

图 4 展示了 GA 和 MCTS 算法下均衡度随试验运行时间的变化规律,带方块黑色虚线和带三角黑色实线分别对应运行 10 次后 GA 和 MCTS 算法的平均均衡度,对应的填充区域为 10 次运行中以最大值和最小值为边界围成的区域,边界线的阶梯下降表明找到了均衡度更小的解。随着试验运行时间的增加,均衡度均值曲线趋于平缓,这是因为要找更优的解已变得困难。相较于 GA,MCTS 算法生成的调度表对应的均衡度曲线下落速率更快,即在相同的试验运行时间内 MCTS 算法能找到更好的解。

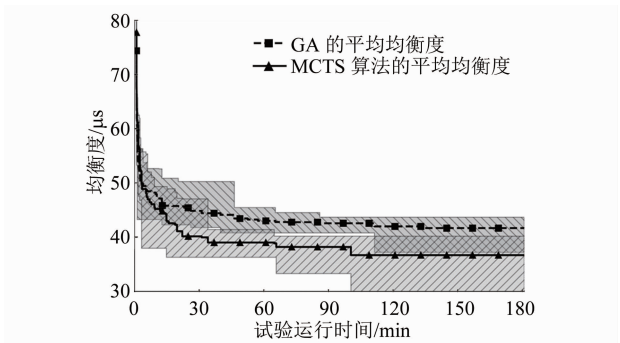


图 4 GA 和 MCTS 算法下均衡度随试验运行时间的变化规律
Fig. 4 Changes in balance score during MCTS and GA iterations

无论是 GA 还是 MCTS 算法,均会使用式(3)衡量生成的每个解的优劣程度。MCTS 算法能统计产生各个解是否比历史解更优等信息,并将这些信息通过回传(步骤 4)存储于祖先节点内,用以指导后续解的生成。也就是说,MCTS 算法中每个新解的生成都参考了历史所有解的信息。相较之下,GA 尽管不断通过优胜劣汰机制来清除劣质解,但每个解的生成仅参考了一个父节点的信息进行变异,而不是全局信息。这也许是 GA 搜索效率不如 MCTS 算法的原因。此外,MCTS 算法还可以通过设置变量优先级来提高算法效果,而 GA 只能随机变异。MCTS 算法的缺点是需要较大内存,用以记录树生成过程中的大量节点。

5 结语

设计负载均衡的 MVB 调度表可减少带宽浪费,并可提高网络通信的实时性。本文首次利用强化学习中的 MCTS 算法实现 MVB 调度表的优化设计,通过设计奖励方式及博弈规则,将 MVB 周期数据调度问题转换成可以用 MCTS 算法解决的组合优化问题。

另外,本文根据 MVB 调度表的特征实现了预剪枝策略,当 MVB 变量总数为 20 个时,可将搜索

空间减少数十个数量级,使其搜索效率优于 GA。仿真结果表明:本文优化后的 MCTS 算法能在相同的搜索时间内获得更均衡的解。如果车辆通信设备增加,需要生成包含更多变量的调度表时,MCTS 算法更能凸显其搜索优势。未来可进一步设计出不同的探索-利用平衡方案,进一步提高 MCTS 算法的搜索效果,以期更快设计出负载均衡的 MVB 调度表。本文给出的设计思路可为通过 MCTS 算法求解其他领域的组合优化问题提供参考。

参考文献

- [1] 谢维达,钱存元. 列车通信与网络[M]. 北京:中国铁道出版社,2014:150.
- XIE Weida, QIAN Cunyuan. Train communication and network [M]. Beijing: China Railway Publishing House, 2014: 150.
- [2] 胡黄水,郑曼,王宏志. 基于遗传算法的 MVB 周期调度表优化[J]. 吉林大学学报(理学版),2019,57(3):613.
- HU Huangshui, ZHENG Man, WANG Hongzhi. Optimization of MVB periodic scheduling table based on genetic algorithm [J]. Journal of Jilin University (Science Edition), 2019, 57(3): 613.
- [3] 王永翔,王立德. 多功能车辆总线周期扫描表的最优化设计[J]. 铁道学报,2009,31(6):46.

WANG Yongxiang, WANG Lide. The optimization method of the MVB period polling table[J]. Journal of the China Railway Society, 2009, 31(6): 46.

- [4] 王涛,王立德,周洁琼. MVB 网络周期轮询算法优化与仿真研究[J]. 机车电传动,2013(6):40.
- WANG Tao, WANG Lide, ZHOU Jieqiong. Cycle polling algorithm optimization and simulation research of MVB network [J]. Electric Drive for Locomotives, 2013(6): 40.
- [5] 朱俊,李芳,王丽芳. 基于蚁群算法的多功能车辆周期扫描表的优化设计[J]. 铁道学报,2013,35(7):57.
- ZHU Jun, LI Fang, WANG Lifang. Optimized design method of periodic list in multifunction vehicle bus based on the ant colony algorithm[J]. Journal of the China Railway Society, 2013, 35(7): 57.
- [6] 张梦璇,于跃,顾宏. 基于改进差分进化算法的 MVB 周期调度表优化设计[J]. 大连理工大学学报,2017,57(2):207.
- ZHANG Mengxuan, YU Yue, GU Hong. Optimized design of MVB periodic dispatch table based on improved differential evolution algorithm[J]. Journal of Dalian University of Technology, 2017, 57(2): 207.
- [7] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676): 354.

(收稿日期:2022-12-17)

(上接第 69 页)

- [4] 高广运,耿建龙,毕俊伟,等. 地铁环境振动对建筑场地影响实测分析[J]. 工程地质学报,2019,27(5):1116.
- GAO Guangyun, GENG Jianlong, BI Junwei, et al. Effect of subway induced environmental vibration on building site through in situ measurement[J]. Journal of Engineering Geology, 2019, 27(5): 1116.
- [5] 李小珍,刘全民,张迅,等. 铁路高架车站车致振动实测与理论分析[J]. 西南交通大学学报,2014,49(4):612.
- LI Xiaozhen, LIU Quanmin, ZHANG Xun, et al. Measurement and theoretical analysis of vehicle-induced vibration on elevated railway station[J]. Journal of Southwest Jiaotong University, 2014, 49(4): 612.

- [6] 刘维宁,张昀青. 轨道结构在移动荷载作用下的周期解析解[J]. 工程力学,2004,21(5):100.
- LIU Weining, ZHANG Yunqing. A periodic analytical solution of railway track structure under moving loads[J]. Engineering Mechanics, 2004, 21(5): 100.
- [7] LOPES P, COSTA P A, FERRAZ M, et al. Numerical modeling of vibrations induced by railway traffic in tunnels: from the source to the nearby buildings[J]. Soil Dynamics and Earthquake Engineering, 2014, 61/62: 269.
- [8] YANG W, WANG M, SHI J, et al. Performance study on the whole vibration process of a museum induced by metro[J]. Structural Engineering and Mechanics, 2015, 55(2): 413.

(收稿日期:2021-05-12)

欢迎订阅《城市轨道交通研究》

服务热线 021—56830728