

基于随机价格时间博弈理论的车-车通信列控系统控制策略稳定性建模与验证

卢万里, 吕继东*

北京交通大学轨道交通运行控制系统国家工程研究中心, 北京 100044

摘要 基于车-车通信的新型列控系统,通过等距离间隔、等时间间隔和变时距3种控制策略实现高效率的列车追踪控制。由于在车队控制中,领航车的控制具有随机性,如何保证不同控制策略中车队的稳定性至关重要。提出了一种基于随机价格时间博弈理论(stochastic priced timed game, SPTG)的车-车通信控制策略建模与验证方法。首先,针对不同控制策略要求,利用随机价格时间自动机,建立包含领航车和跟随车的车队控制模型,并进行稳定性验证;然后,以时间为成本函数,通过对建立车队随机价格时间博弈自动机模型,利用Q-learning强化学习方法得到车队的最优驾驶策略;最后,结合多车运行追踪场景,进行车队的稳定性仿真优化。结果表明:相比于车队的随机运行策略,该方法使得车队的稳定误差更小。

关键词 车-车通信;列控系统;随机价格时间博弈理论;稳定性;建模验证

传统列控系统采用固定闭塞、准移动闭塞和移动闭塞的闭塞方式^[1]进行防护,列车通过与地面设备、车站的通信获得线路信息及线路上列车信息。由于这种传统方式的技术和规定限制,使得线路上列车的行车密度始终无法达到合理的分布和线路的最大运载,浪费线路的时空资源。

近年来,随着通信技术的发展,基于车-车通信的新型列控系统得到了广泛关注^[2-4]。针对列控

系统普遍存在的问题,国内外学者结合无线通信、卫星定位、自动控制等信息技术在列控系统的应用,开展了广泛的新型列车控制系统研究^[5-9]。

围绕实现行车安全、提高线路运行效率及车速等优化目标,基于车-车通信的新型列控系统可采用固定追踪空间间隔、固定追踪时间间隔、可变追踪间隔(空间和时间)等不同的控制策略灵活地控制列车。通过将线路上的列车编组为车队,车队中

收稿日期:2022-12-08;修回日期:2023-02-16

基金项目:国能铁路装备有限责任公司先进轨道交通综合试验研究基地方案研究项目(TZKY-21-16);国家自然科学基金项目(52272329);北京市自然科学基金项目(L211019, L201004);中国国家铁路集团有限公司科技研究开发计划项目(L2021G003)

作者简介:卢万里,博士研究生,研究方向为列车控制系统,电子信箱:luwanli@bjtu.edu.cn;吕继东(通信作者),教授,研究方向为轨道交通,电子信箱:jdlv@bjtu.edu.cn

引用格式:卢万里,吕继东. 基于随机价格时间博弈理论的车-车通信列控系统控制策略稳定性建模与验证[J]. 科技导报, 2023, 41(10): 82-91; doi: 10.3981/j.issn.1000-7857.2023.10.007

跟随列车以前车车速和位置信息作为本车控车依据(领航车以追踪目标点为控车依据)。由于领航车的加速和制动的动作是随机进行的,影响到车队中列车的运行,跟随列车在控车过程中受到前车运行波动的影响,所以,车队在线路上运行的过程中找出领航车的最优控制策略以消减车队中跟随列车的速度位置波动、保证车队运行稳定性及提高行车密度至关重要。

博弈论(game theory)^[10-11]是多局中人获取最优解的一种优良方法,本研究车队中多列车共同参与决策,达成有约束力的协议,可采用合作博弈^[12]的方式。形式化方法近年来在优化控制方面被广泛应用,在汽车、铁路、航天中均取得了许多优良成果^[13-15]。

随机价格时间博弈(stochastic priced timed game, SPTG)^[16]理论在求解多对象合作博弈最优解中效果优良,其通常在保证环境随机的前提下,找到一种能最大程度减少预期可达性成本的策略。在此基础上本研究将结合强化学习^[11]中的Q-learning^[17]方法,通过观察博弈产生的所有变迁轨迹,从中学习并迭代地改进控制策略以达到相对更优的控制策略结果。

此外形式化方法基于严格的数学定义,在系统的安全性、可达性等方面的验证中更加严格有效^[18]。

车辆跟驰控制一直以来是交通界非常关注的问题,文献[19]提出了一种改进的多车跟随(multiple car-following, MCF)模型,利用线性稳定性理论获得模型线性稳定性条件,通过非线性分析得出改进的mKdV(Korteweg-de Vries)方程来描述临界点附近交通行为。文献[20]中采用多辆汽车车队的跟车实验,研究各种交通条件下汽车速度与其间距之间的关系。在铁路领域,列车间的通信可以更好地保证行车的安全性。文献[21]中选取两车区间追踪场景,通过CPN(colored petri net)对车-车通信列控系统顶层进行建模与仿真,验证了系统的性质及功能。列车控制重点关注列车运行的稳定性以及行车间距的合理性。文献[22]首先介绍在多列列车中的协调控制方法和相应的稳定性判据,

之后考虑了2种协调方案并提出相应的控制算法,最后利用Lyapunov和不变集定理建立了它们的稳定性。

本研究采用基于SPTG理论的控制策略及控制方法,针对新型列控车-车通信控制策略进行建模与验证。首先,设计列车车队中领航车与跟随车的控制方程,利用随机价格时间自动机,建立包含领航车和跟随车的车队控制模型,针对不同控制策略要求,进行稳定性证明;其次,以时间为成本函数,通过对车队随机价格时间自动机模型的合作博弈,利用Q-learning得到车队的最优驾驶策略;最后,结合多车运行追踪场景,进行车队的稳定性仿真优化。

1 新型列车控制系统的控制策略

1.1 新型列控系统的控制策略提取

将列车车队建模为八元组 $(x, n, m, k, c, h, c_h, v_d)$,其中每个参数的描述、类型和约束见表1。

表1 车队参数介绍

参数	类型	要求
水平距离 x	R	$x > 0$
车辆数量 n	N	$n > 1$
车辆质量 m	R	$m > 0$
干扰常数 k	R	$k > 0$
波动常数 c	R	$c > 0$
时间差 h	R	$h > 0$
时间波动导致的波动 c_h	R	$c_h > 0$
车队所需速度 v_d	R	$v_d > 0$

稳定性是车队的重要属性,描述了车队衰减由车队队首车辆引入振荡的能力。可以根据各种信号来考虑这种振荡,例如,列车间的位置误差或列车的相对加速度。形式上,如果在朝向队尾传播时,相对于特征误差的任何振荡减小,则车队是稳定的。

列车车队稳定性在数学上表示为频域中间距误差的标准条件,计算公式为

$$\left| \frac{z_i}{z_{i-1}} \right| \leq 1 \quad (1)$$

式中, z_i 为列车 i 与其前行列车 $i+1$ 之间的间隔误差, $z_i = x_i - x_{i+1}$ 。在所有期望的列车间距都相同的情况下, 即 $x_i = x_{i-1} = \dots = x_1$, 那么这将得到零间隔误差, 即 $z_i = z_{i-1} = \dots = z_1 = 0$ 。

1.2 控制方程构建

车队中领航车考虑追踪点的位置, 而跟随车的控制主要参考前车的位置和速度, 故构建列车控制方程

$$\begin{cases} \frac{dx_i}{dt} = v_i, & i = 1, 2, \dots, n \\ \frac{dv_i}{dt} = \frac{u}{m}, & i = 1 \\ \frac{dv_i}{dt} = j_{i-1}x_{i-1} - f_i x_i + j_{i-1}v_{i-1} - f_i v_i, & i = 2, 3, \dots, n \end{cases} \quad (2)$$

式中, x, v 分别为车队的车辆间距和速度; j, f 分别为位置波动与列车质量的比值、速度波动与质量的比值; u 为列车牵引力, 稳定性可以通过消除变量 v 并将其以间距误差的形式表示。

取 $j_{i-1} = j_i, f_{i-1} = f_i$, 通过消去参数, 其稳定性公式为

$$\frac{d^2 z_i}{dt^2} + f_i \frac{dz_i}{dt} + j_i z_i = f_i \frac{dz_{i-1}}{dt} + j_i z_{i-1} \quad (3)$$

通过拉普拉斯变换, 可将式(3)转入复频域分析, 转换后的对应象函数方程为

$$\frac{z_i(s)}{z_{i-1}(s)} = H(s) = \frac{f_i s + j_i}{s^2 + f_i s + j_i} \quad (4)$$

式(4)显示的控制方式为带比例微分的二阶系统。通过与式(5)传递函数对比该式显示的控制方式为带比例微分的二阶系统。

$$H(s) = \frac{T_d \omega_i^2 s + \omega_i^2}{s^2 + (2\xi \omega_i + T_d \omega_i^2) s + \omega_i^2} \quad (5)$$

在列车波动质量比相同情况下, f_i 对应 $T_d \omega_i^2$ 参数, j_i 对应 ω_i^2 参数。等效阻尼比为 $\frac{f_i}{(2\sqrt{j_i})}$, 振荡倾向和超调量减小, 在实际超调出现之前就产生一个适当修正, 使得系统既有较高的稳态精度, 又有良好的平稳性。所以列车类型与质量相同时, 列车间隔误差收敛, 车队运行能够趋于稳定。

1.3 控制策略构建

结合不同控制策略的需求特性, 构建控制方程的过程中应考虑不同的参数, 对应参数如表2所示。

表2 各控制策略所需参数

控制策略	j_i	f_i
固定时间间隔控制策略	$\frac{k}{m}$	$\frac{c + kh}{m}$
固定空间间隔控制策略	$\frac{k}{m}$	$\frac{c}{m}$
变时距控制策略	$\frac{k}{m}$	$\frac{c + kc_k v_d}{m}$

2 SPTG 理论优化

2.1 SPTG

价格时间博弈^[23]是在价格时间自动机^[24]上进行博弈。价格时间博弈自动机是八元组, 表示为 $K = (L, \ell_0, X, \Sigma_c, \Sigma_u, E, P, I)$ 是八元组, 其中: L 为位置状态的有限集; $\ell_0 \in L$ 为初始状态位置; X 为一组实数值时钟; Σ_c 为可控动作集合; Σ_u 为不可控动作集合; E 为变迁的有限集; P 为每个状态分配价格率; I 为每个状态的时钟值域。

SPTG K 语义 $S_c = (Q, q_0, \Sigma, \rightarrow)$ 是价格变迁系统, 其中状态 Q 由 (ℓ, κ) 对组成, $\ell \in L$ 且 $\kappa \in \mathbb{R}_{\geq 0}^X$ 时 $\kappa \models I(\ell)$, 其中 $\mathbb{R}_{\geq 0}^X$ 为 X 维大于 0 的实数空间, $q_0 = (\ell_0, 0)$ 为初始状态。 $\Sigma = \Sigma_c \cup \Sigma_u$ 为动作的有限集, $\rightarrow \subseteq Q \times (\Sigma \cup \mathbb{R}_{\geq 0}) \times \mathbb{R}_{\geq 0} \times Q$ 为动作 $a \in \Sigma$ 和延迟 $d \in \mathbb{R}_{\geq 0}$ 分别定义的价格变迁关系, 动作变迁的价格为 0, 而延迟变迁的价格则根据给定位置的价格比率与延迟成比例。

SPTG K 的 run 是其价格变迁系统 S_K 的价格动作和延迟转换的交替序列:

$$\pi = q_0 \xrightarrow{d_0} p_0 q'_0 \xrightarrow{a_0} q_1 \xrightarrow{d_1} p_1 q'_1 \xrightarrow{a_1} \dots \xrightarrow{d_{p-1}} p_{p-1} q'_{p-1} \xrightarrow{a_{p-1}} q_p \dots \quad (6)$$

式中, $a_p \in \Sigma$; $d_p, p_p \in \mathbb{R}_{\geq 0}$; q_i 为状态 (ℓ_p, κ_p) 。 K 的 run 集合表示为 $\Pi_K, \Pi'_K (\Pi''_K)$ 为其有限(最大)run 的集合。对于有限 run $\pi, C(\pi)$ 为其总累积成本 $\sum_{p=1}^N p_p$,

$T(\pi)$ 为总累积时间 $\sum_{\rho=1}^N d_\rho$, $\text{last}(\pi)$ 表示 π 的最后状态。

策略是一种功能,在博弈过程中不断为控制器提供有关为赢得博弈而应该采取的措施信息。令 $K \subseteq L$ 为目标位置集, $B \in \mathbb{R}_{\geq 0}$ 为成本(时间)的限制, K 的 (K, B) 受成本(时间)限制的可达性控制问题在于找到一个强制 K 在累积成本(时间) B 内到达的策略 σ 。

令 (G, B) 为 K 上的可达性博弈。 π 如果存在 $\text{last}(\pi) \in G \times \mathbb{R}_{\geq 0}^x$ 且 $C(\pi) \leq B$, 则为获胜 run, 获胜 run 集合表示为 Π_w 。结果 Π_o 是 Π_k 的子集, 如果 Π_o 中的所有 run 都在 Π_w 中, 则策略 σ 是获胜策略。

SPTG 中, 延迟和动作的选择均是随机的, 服从于给定的概率密度函数, 本研究选用均匀分布进行刻画。状态 q_ρ 处延迟约束范围的 $[d_{\min}, d_{\max}]$ 概率密度分布为

$$\mu_\rho(d_\rho, a_\rho) = \frac{d_\rho - d_{\min}}{d_{\max} - d_{\min}} \quad (7)$$

式中, $\mu_\rho(d_\rho, a_\rho)$ 为概率密度分布函数; d_{\min} 为最小时延; d_{\max} 为最大时延。

2.2 SPTG 优化

价格时间马尔可夫决策过程 (priced timed markov decision process, PTMDP) 为 $M=(G, \mu_u, \mu_c)$, 其中 $\mu_u = \{\mu_\rho(d_\rho, a_\rho) | a_\rho \in \Sigma_u\}$ 为从状态 q_ρ 延迟 d_ρ 后发生控制器不可控制动作的概率密度分布族; $\mu_c = \{\mu_c(d_\rho, a_\rho) | a_\rho \in \Sigma_c\}$ 为从状态 q_ρ 延迟 d_ρ 后控制器选择可控动作的概率密度分布族。 M 和 μ^c 的组合定义了 run 集上的概率测度 \mathbb{P}_{M, μ^c} 。

本研究旨在得到接近最优的控制策略 μ^o 来最小化成本(时间)期望值, 以 $\sigma^o(G, T)$ 为 M 包含所有可控动作的有时间限制的可达性策略, 即最宽松的控制策略。

令 $\pi \in \Pi_M$, G 作为目标位置集合, $C_c(\pi) = \min\{C(\pi_i) : \ell_i \in G\}$ 表示 π 到达 G 之前的累积成本。 C_c 为一个随机变量, 对于给定随机策略 μ^c , 期望值 $\mathbb{E}_T^M(C_c) = \int_{\pi \in \Pi_M} C_c(\pi) \mathbb{P}_{M, \mu^c}(d\pi)$ 。

给定 PTMDP M 和随机控制策略 μ^c 生成许多 run, 这些 run 既用于估计 $\mathbb{E}_T^M(C_c)$, 又用于朝 μ^o 迭代地改善 μ^c 。算法上即要保证给定的时间限制, 又要实现预期成本的最优性。

图 1 为本文算法的一般流程。

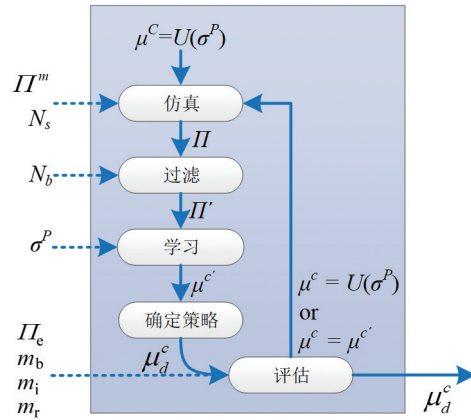


图 1 使用强化学习的最佳调度器近似

具体步骤为:(1) 生成最多用于学习的 run Π_M^m , 最多生成 N_s 次。结果是一组 run 的集合 Π 。 Π 可能为空, 在这种情况下, μ^c 会重置, 仿真会重新开始;(2) 对 run Π 集合进行过滤, 并保留其中最多的 N_b 最佳 run。保留成本最低的 run $\Pi' \subseteq \Pi$ 子集;(3) 进入学习阶段, 采用 Q-learning 强化学习方法计算出新的 μ^c' ;(4) 在评估之前确定最终的策略。此步骤再次评估变迁路径 Π_c 上的强化策略 μ_d^c 的成本, 在生成的 μ^c' 的成本低于 μ^c 的情况下, 更新 μ^c 。但是, 如果 μ^c' 不超过 μ^c 的 m_b 倍, 则将重置为 $\mu^c = U(\sigma^P)$, 允许所有可能的动作仍然保证控制器在给定的时间内达到其目标, 确保强化学习不会陷入局部最小值。(5) 如果重置 μ^c 不超过 m_i 次, 即算法最多循环 m_i 次。

当算法停止时, 将输出最好的确定性策略 μ_d^c 。

2.3 Q-learning

Q-learning 强化学习方法具有计算智能体状态-动作组合 (q, a) 价值的功能, 包含了智能体、环境状态 q 、奖励 r 、动作 a , 已学习经验矩阵 Q , 收益矩阵 R 。转移规则为

$$Q(q,a) = R(q,a) + \gamma \max_{\tilde{a}} Q(\tilde{q},\tilde{a}) \quad (8)$$

式中, γ 为奖励性衰变系数; (\tilde{q},\tilde{a}) 为 (q,a) 的下一个状态。算法如表3所示。

表3 Q-learning算法

Q-learning算法	
Step 1	给定参数 γ 和reward矩阵 R ;
Step 2	令 $Q := 0$;
Step 3	对每一轮循环
1.	随机选择一个初始状态 q ;
2.	若未达到以下状态,则执行以下步骤:
(1)	在当前状态 q 的所有可能行为中选取一个行为 a ;
(2)	利用选定的动作 a ,得到下一个状态 \tilde{q} ;
(3)	根据式(1)计算 $Q(q,a)$;
(4)	令 $q := \tilde{q}$ 。

2.4 UPPAAL-Stratego

采用UPPAAL-Stratego^[25]进行车-车通信控制模型各个模块的构建,根据时间博弈自动机的积组成时间博弈自动机网络模型。UPPAAL-Stratego结合符号综合和强化学习,可以提供随机时间博弈合成安全和接近最优的策略。该工具最近已扩展到随机混合博弈,并成功应用于电池感知调度问题的策略合成^[26],以及汽车的安全性和最佳自适应巡航控制器^[27],如图2所示。

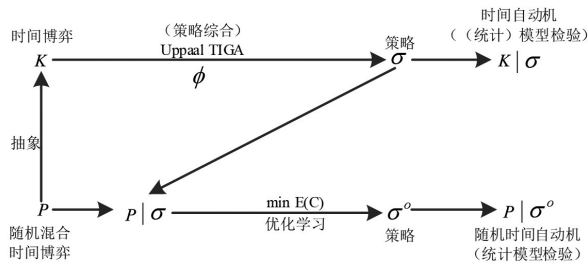


图2 Uppaal-Stratego概述

在控制策略合成与优化过程中,首先,通过忽略模型中的价格和随机性,UPPAAL-Stratego将SPTG P 抽象为时间博弈(TGA) K ,使用UPPAAL-TIGA合成满足所需安全性或(有时间限制)活动约束 ϕ 的策略 σ ;然后,使用强化学习合成表示成本期望值最大化的近似最优策略 σ^o ;最后,在给定 σ^o 的情况下,可以执行对 $P | \sigma^o$ 的统计模型检验。

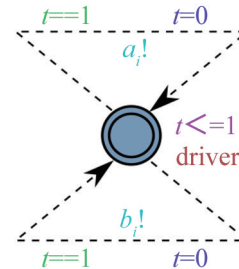
3 车-车通信列车控制模型

以区间追踪场景中列车运行路径内有多列车情况进行列车运行策略的构建,不考虑地面对象控制器(object controller, OC)的存在。

基于对车-车通信控制的考虑,本文模型关注于列车状态的仿真,对于地面设备默认都是正常且反应正确的。

3.1 列车动力学模型

列车动力学模型为列车提供随机加速或减速动作。模型如图3所示,模型仿真车载计算机在控车过程中随机指令的下达。



b_i 表示下达减速指令, a_i 表示下达加速指令; t 为周期

图3 列车动力学模型

3.2 列车速度控制模型

列车速度控制模型如图4所示,主要为加减速Acc与Dec这2个状态、列车惰性状态Maintain,以车队跟随者角色运行时的Follower状态。

列车在车队中为领航者角色或以单车运行时,模型周期输出列车的运行速度、并计算行车许可(movement authority, MA)及列车目标速度。列车的输出牵引力或制动力公式为

$$\frac{dv_i}{dt} = (v_{0,i})F - F_{r,i} \quad (9)$$

式中, v_i 为列车当前速度, $v_{0,i}$ 理论上列车当前运行速度, F 为动力参数, $F_{r,i}$ 为线路上的阻力(包括钢轨摩擦、坡度、空气阻力等)与列车质量的比值。

列车以跟随者角色运行时可采用惰性状态Maintain及Follower状态进行制动或牵引。其Follower状态通过采集前车距追踪点的距离、本车距前车车尾追踪点的距离和前车与本车的速度来进

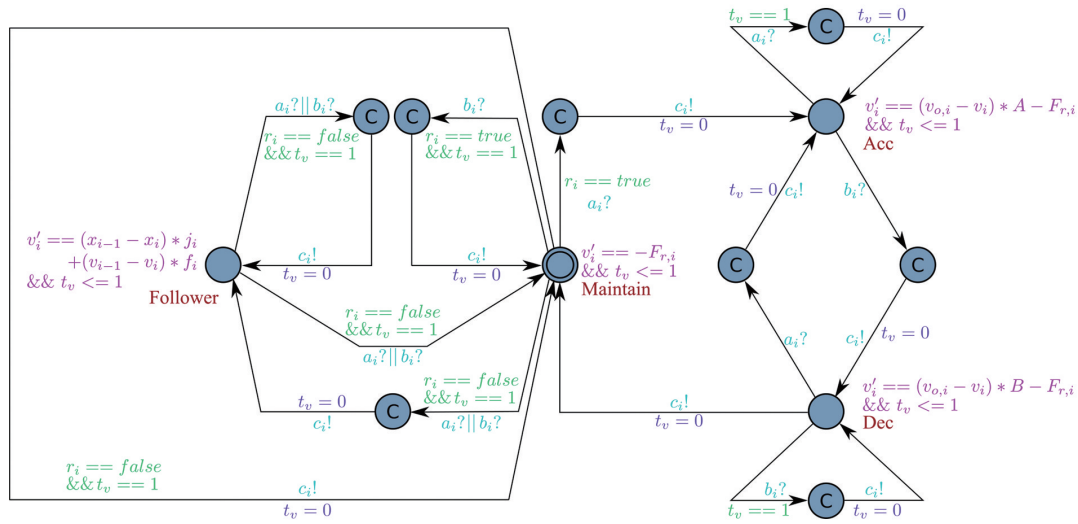


图4 列车速度控制模型

行列车的速度控制。其方程如式(10)所示。

$$\frac{dv_i}{dt} = j_i * (x_{i-1} - x_i) + f_i * (v_{i-1} - v_i) \quad (10)$$

列车在运行过程中从惰性状态出发,依次进行列车速度控制、MA更新和目标速度更新。模型变量介绍见表4。

表4 列车速度控制模型变量介绍

变量	类型	含义
A	double array	列车加速过程动力参数
B	double array	列车制动过程制动参数
t_v	clock	列车速度控制模型运行参数
c_i	broadcast chan	计算列车距追踪点距离指令
r_i	bool array	列车在车队中的角色

3.3 列车目标速度和移动授权计算以及交权模型

MA在计算时分为领航者模式和跟随者模式。当前列车为领航者角色时,其在列车速度大于零且距离追踪点距离大于零的情况下工作,计算公式为

$$x_i = x_i - \frac{v_i}{1000} t \quad (11)$$

当列车在制动范围之外时,其目标速度设定为其理论计算运行速度。列车在达到制动范围时,其限速计算如式(12)所示。

$$\frac{dv_{o,i}}{dt} = -\frac{v_{o,i}^2}{(2000x_i)} + \lambda \quad (12)$$

式中, $v_{o,i}$ 为列车理论目标速度,由于周期均以 s 为单位,故速度均取 m/s 为单位。

由于距离单位取 m 会导致数字过大,为便于书写和使用,选定距离单位为 km ,当运算时再进行单位的换算,自动机计算过程中会丢失部分结果,导致计算失真,故添加补充变量 λ 。当列车到达追踪点后,其将车队的控制权交给其后方的第一辆车。

当前列车为跟随者角色时,其在计算与前行追踪点的距离过程中需考虑前车的当前运行速度,并将其纳入计算模型中,计算公式为

$$x_i = x_i - \frac{v_i}{1000} t + \frac{v_{i-1}}{1000} t \quad (13)$$

跟随者角色的列车在制动距离外与领航者列车的运行方式相同,但当列车与前行列车车尾间距小于等于制动距离时,以前行列车的当前运行速度为目标速度来刻画本车的运行速度,车间距的变化率为本车与前车的速度之差与计算周期的比值,即

$$\frac{dv_{o,i}}{dt} = -\frac{(v_i - v_{i-1})}{t} \quad (14)$$

建立模型如图5所示, δ 为列车制动距离。

在列车运行过程中,控制列车当前的特征变量更加快速达到目标值的控制策略是最优的。通过以误差差值最小为控制目标进行控制策略的学习和迭代,获得最优控制策略。图6为控制策略目标

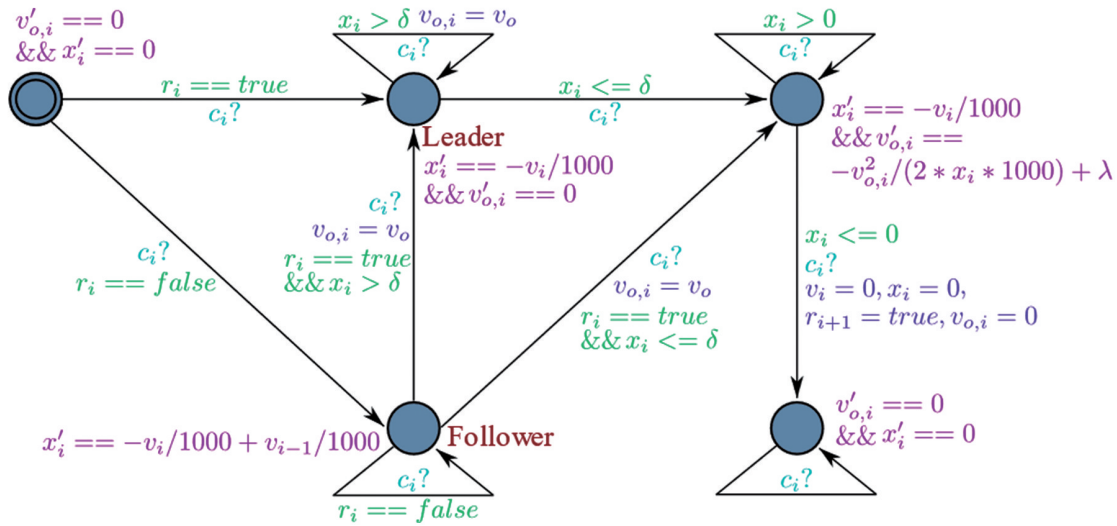


图5 列车目标速度和距离追踪点距离计算以及交权模型

值计算过程,图中 $D_i(t)$ 表示列车 i 的控制策略目标值。

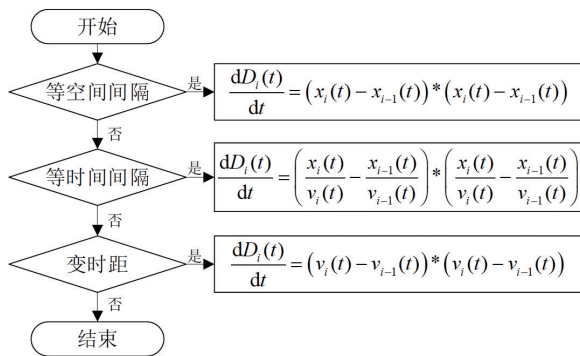


图6 控制策略目标值计算

4 车-车通信控制策略优化分析

以车队中列车质量相同的下的不同控制策略进行介绍。

4.1 等空间间隔控制策略

若线路上列车均相同,则可监控其运行速度或运行间隔来进行车队的控制。

以下以车队内有4列车为例进行验证。在本策略运行下的车队其相对运行速度相差不大,保证了列车运行安全性。车队中队首列车采用领航者控制策略,第2列车的车距参照制动距离,其后列

车均参照前行列车与前行列车间的间隔,主要研究列车运行稳定性,所以此处设定各列车的起始速度均为最大速度,以便观察稳定性特征。随机运行仿真图与基于策略的仿真图如图7(a)、图7(b)所示。

本策略采用监控车队运行间隔的方式进行车队运行控制。该要求保证列车与列车间的间隔保持相对稳定,其稳定性可按式(1)表达。在实际运算中, z_i 会遇到为零的情况,故对式(1)进行变形,即

$$-z_{i-1} \leq z_i \leq z_{i-1} \quad (15)$$

式(15)表示 i 车与 $i-1$ 车间距的差值在 $i-1$ 车和 $i-2$ 车间距的差值正负值之间。

根据图7(b)所示,列车车队领航者大约在 $t=500$ s 左右时进行制动,故分析 500 s 内的控制策略稳定性。得到的随机运行结果与基于策略的运行结果如图8所示,显示了在 500 s 内列车运行的差值,可以看到相对于随机运行策略,在基于最优综合控制策略下误差更小,且 z_3 更能包络在 z_2 的正负值之间,振幅更小。说明该策略对达到稳定性要求更具有效果。

4.2 等时间间隔控制策略

本研究主要以等间隔控制策略阐述获取最优控制策略的思想,与其余控制策略优化的核心思想是一致的,故以下对这2种控制策略列车速度位置曲线以及相应稳定性参考依据的仿真图简略表述。

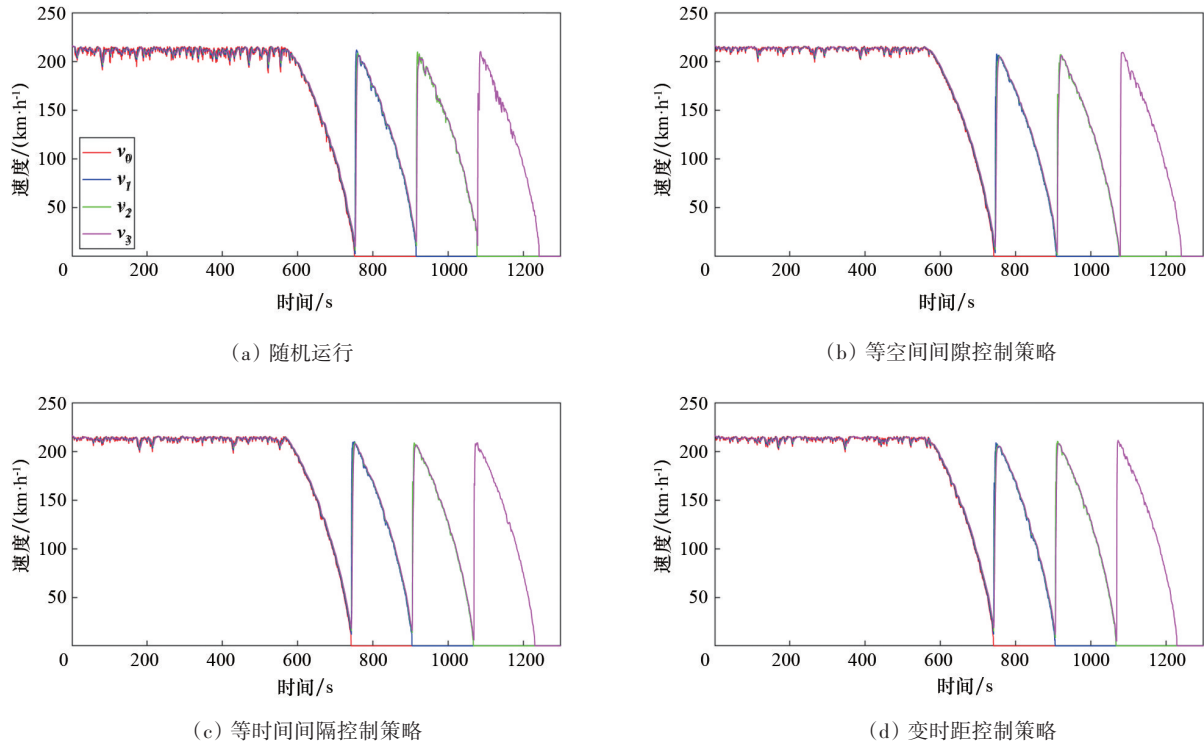


图7 四列车运行速度仿真

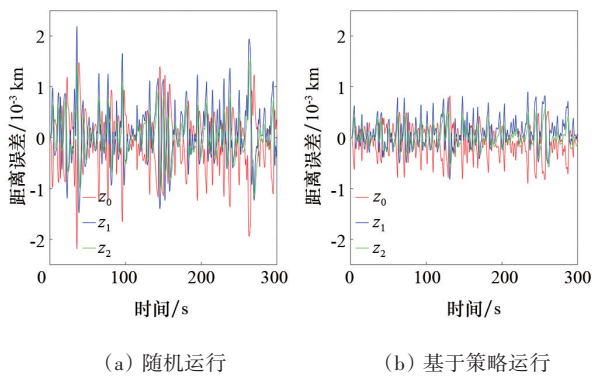


图8 等空间间隔控制策略稳定性

等时间间隔控制策略下列车的当前速度与车距的比值是基本稳定的,故取判据进行稳定性判断。计算公式为

$$\begin{cases} t_i = x_i / v_i \\ z_i = t_i - t_{i-1} \Rightarrow z_i = \frac{x_i v_{i+1} - x_{i-1} v_i}{v_i v_{i+1}} \end{cases} \quad (16)$$

由于在运算过程中分母可能为零导致运算错误,为避免错误,可在优化过程中尽量使得分子为

最小且分母尽量最大。以此进行仿真分析。仿真图如图7(c)和图9所示,可看出相较随机运行,基于策略运行的情况下时间误差振幅更小,更准确。

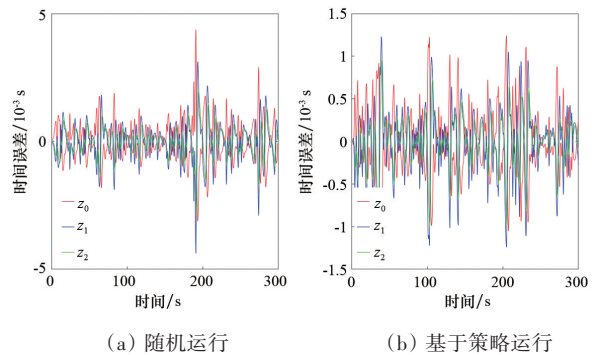


图9 等时间间隔控制策略稳定性

4.3 变时距控制策略

变时距控制策略考虑“撞软墙”形式,列车的速度保证与前车一致即可,故稳定性判断可以对两车速度进行判断,即

$$z_i = v_i - v_{i-1} \quad (17)$$

以式(17)进行仿真判断,结果如图7(d)和图10所示,在基于策略运行情况下,速度误差振幅降低了50%。

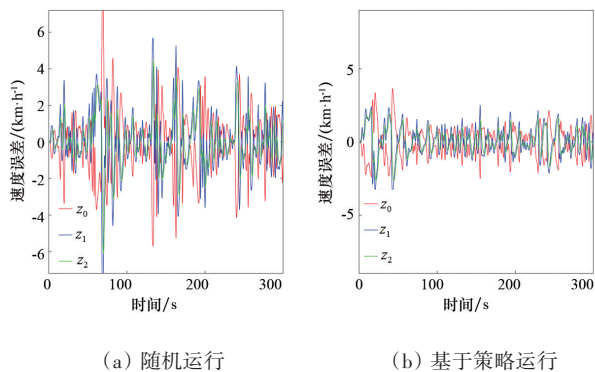


图10 变时距控制策略稳定性

5 结论

提出了一种列车车队的控制方式,采用控制方程进行描述,并根据不同控制目标要求进行了稳定性证明。通过采用SPTG理论,对于提出的控制方程进行建模,针对所提出的3种车队控制策略进行了证明和验证,其结果表明这些控制策略能够实时控制线路上车队的运行,在保证稳定性的前提下间距误差更小,有效提高了线路的利用率,为未来基于车-车通信新型列控系统提供了理论参考。本文所提出的方法主要解决了线路上列车车队运行稳定性和效率提高问题,为车-车通信下的车队运行控制提供了多样化的控制策略。此外,将进一步研究基于SPTG理论的控制策略及控制方法在不同运行场景下的应用,研究采用多种控制策略下的行车效率问题,保证控制策略的可靠性。

参考文献(References)

- [1] 郜春海. 基于通信的列车运行控制(CBTC)系统[M]. 北京: 中国铁道出版社, 2018: 40-95.
- [2] Zhu L, Yao D Y, Zhao H L. Reliability analysis of next-generation CBTC data communication systems[J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 2024-2034.
- [3] Chen K H, Lv J D, Luo Z W, et al. Complete testing for speed monitoring function of next-generation train control system based on IPOG strategy[C]//2019 IEEE Intelligent Transportation Systems Conference (ITSC). Piscataway: IEEE Press, 2019: 3633-3638.
- [4] Yu F R. Advances in communications-based train control systems[M]. New York: CRC Press, 2015.
- [5] 赵磊, 何春明. 美国PTC系统和欧洲ERTMS的差异分析[J]. 铁道通信信号, 2011, 47(11): 56-59.
- [6] Lindsey R. Positive train control in North America[J]. IEEE Vehicular Technology Magazine, 2009, 4(4): 22-26.
- [7] Lei L, Lu J H, Jiang Y M, et al. Stochastic delay analysis for train control services in next-generation high-speed railway communications system[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(1): 48-64.
- [8] Raupp G, Behler K, Cole R, et al. Next generation discharge control system for ASDEX upgrade[J]. Fusion Engineering and Design, 1999, 46(2-4): 347-354.
- [9] Gurník P. Next generation train control (NGTC): More effective railways through the convergence of main-line and urban train control systems[J]. Transportation Research Procedia, 2016, 14: 1855-1864.
- [10] Hai X S, Wang Z L, Feng Q, et al. A novel adaptive pigeon-inspired optimization algorithm based on evolutionary game theory[J]. Science China Information Sciences, 2021, 64(3): 139203.
- [11] Abdoos M. A cooperative multiagent system for traffic signal control using game theory and reinforcement learning[J]. IEEE Intelligent Transportation Systems Magazine, 2021, 13(4): 6-16.
- [12] 姜启源, 谢金星, 叶俊. 数学模型[M]. 4版. 北京: 高等教育出版社, 2011: 373-410.
- [13] Rashid A, Siddique U, Hasan O. Formal verification of platoon control strategies[C]//Johnsen E, Schaefer I. International Conference on Software Engineering and Formal Methods. Cham: Springer, 2018: 223-238.
- [14] 吕继东. 列车运行控制系统分层形式化建模与验证分析[D]. 北京: 北京交通大学, 2011.
- [15] 张锦坤, 杨孟飞, 乔磊, 等. 基于有限状态机的操作系统需求层形式化验证[J]. 空间控制技术与应用, 2019, 45(2): 48-55.
- [16] David A, Jensen P G, Larsen K G, et al. On time with minimal expected cost! [M]//Automated Technology for Verification and Analysis. Cham: Springer International Publishing, 2014: 129-145.
- [17] Christopher J C H. Q-learning [J]. Machine Learning,

- 1992, 8(3-4): 279-292.
- [18] 曹源, 唐涛, 徐田华, 等. 形式化方法在列车运行控制系统中的应用[J]. 交通运输工程学报, 2010, 10(1): 112-126.
- [19] Peng G H, Sun D H. A dynamical model of car-following with the consideration of the multiple information of preceding cars[J]. Physics Letters A, 2010, 374(15-16): 1694-1698.
- [20] Jiang R, Hu M B, Zhang H M, et al. On some experimental features of car-following behavior and how to model them[J]. Transportation Research Part B: Methodological, 2015, 80: 338-354.
- [21] 王鹏, 李开成, 刘雨. 车车通信技术在列控系统中的应用研究[J]. 铁道通信信号, 2016, 52(7): 62-65, 70.
- [22] Dong H R, Gao S G, Ning B. Cooperative control synthesis and stability analysis of multiple trains under moving signaling systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2730-2738.
- [23] Alur R, Dill D L. A theory of timed automata[J]. Theoretical Computer Science, 1994, 126(2): 183-235.
- [24] Behrmann G, Fehnker A, Hune T, et al. Minimum-cost reachability for priced time automata[M]//Hybrid Systems: Computation and Control. Berlin, Heidelberg: Springer, 2001: 147-161.
- [25] David A, Jensen P G, Larsen K G, et al. Uppaal stratego [C]//Baier C, Tinelli C. TACAS 2015: Tools and Algorithms for the Construction and Analysis of Systems. Berlin: Springer, 2015: 206-211.
- [26] Wognsen E R, Haverkort B R, Jongerden M, et al. A score function for optimizing the cycle-life of battery-powered embedded systems[C]//International Conference on Formal Modeling and Analysis of Timed Systems. Berlin: Springer, 2015: 305-320.
- [27] Larsen K G, Mikuionis M, Taankvist J H. Safe and optimal adaptive cruise control[M]//Meyer R, Platzer A, Wehrheim H. Correct System Design. Cham: Springer, 2015: 260-277.

Modeling and verification of stability control strategy of train-train communication train control system based on SPTG

LU Wanli, LV Jidong*

National Engineering Research Center of RailTransportation Operation and Control System, Beijing Jiaotong University Beijing 100044, China

Abstract Next generation train control system (NGTC) based on Train-Train communication realizes highly efficient train tracking control through three control strategies of constant spacing, constant time interval, and dynamic headway. Since the leader train control is random in the platoon control, how to ensure the safety of the platoon in different control strategies is very important. This paper proposes a Train-Train communication control strategy modeling and verification method based on stochastic priced timed game (SPTG). Firstly, according to the requirements of different control strategies, a platoon control model including a leader train and follower trains is established by using SPTG automata, and the stability is verified. Secondly, taking time as the cost function, Q-learning is used to obtain the optimal driving strategy of the platoon through the platoon's SPTG automata model. Finally, combined with multi-train operation tracking scenarios, the stability simulation optimization of the platoon is carried out. The result shows that the stability error of the platoon is smaller than that of the random operation of the platoon.

Keywords train-to-train communication; train control system; SPTG; stability; modeling and verification ●



(责任编辑 刘志远)