



中国科技期刊卓越行动计划项目入选期刊

控制与决策

CONTROL AND DECISION

车路协同环境下数据驱动的混合交通流速度调控方法

张辰, 徐云雯, 李德伟

引用本文:

张辰, 徐云雯, 李德伟. 车路协同环境下数据驱动的混合交通流速度调控方法[J]. *控制与决策*, 2024, 39(9): 2950–2958.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0700>

您可能感兴趣的其他文章

Articles you may be interested in

[基于混合整数规划的智能网联车冲突区时序优化模型](#)

Mixed integer programming model of scheduling for connected automated vehicles in a conflict zone

控制与决策. 2021, 36(3): 705–710 <https://doi.org/10.13195/j.kzyjc.2019.0886>

[区分交通流模式的混合服务路口信号控制策略](#)

Signal control strategies of mixed service intersections to discriminate traffic flow patterns

控制与决策. 2021, 36(6): 1509–1515 <https://doi.org/10.13195/j.kzyjc.2019.1520>

[基于MCPDDPG的智能车辆路径规划方法及应用](#)

The method and application of intelligent vehicle path planning based on MCPDDPG

控制与决策. 2021, 36(4): 835–846 <https://doi.org/10.13195/j.kzyjc.2019.0460>

[车辆跟随控制策略的状态可达集建模及验证方法](#)

A modeling and verification method of state reachable set for vehicle following control strategy

控制与决策. 2021, 36(7): 1679–1685 <https://doi.org/10.13195/j.kzyjc.2019.1562>

[基于强化学习的多目标车辆跟随决策算法](#)

Multi-objective vehicle following decision algorithm based on reinforcement learning

控制与决策. 2021, 36(10): 2497–2503 <https://doi.org/10.13195/j.kzyjc.2020.0426>

车路协同环境下数据驱动混合交通流速度调控方法

张辰, 徐云雯[†], 李德伟

(上海交通大学 自动化系, 上海 200240)

摘要: 针对车联网和自动驾驶背景下的混合车流调控问题, 提出通过路侧设备调节车流中受控自动网联车速度以实现宏观交通流速度调控的方法. 首先将单个受控自动网联车对车流状态的调节过程描述为马尔可夫决策过程, 设计衡量受控车辆对整体车流影响程度的动态网格, 以受控车辆的实时期望速度为动作, 并综合考虑受控车对网格车流的影响程度、对期望速度的响应快慢以及车流的安全性构造报酬函数, 采用深度确定性策略梯度算法进行策略寻优. 考虑到受控自动车周边交通环境的多样性, 提出基于关键参数量的控制策略模型簇. 其次, 面向车流速度的实时调控, 根据路段车辆簇的划分, 提出多个路侧设备控制同一车辆簇的控制权时序交接策略. 路侧设备以簇为单位对信号覆盖范围内的车辆进行标记与跟踪, 实时计算受控车辆的期望控制信号并下发. 最后, 将所提出的方法在多种场景下进行训练验证, 结果表明所提出的方法在适用场景中能够实现车流的平稳调速.

关键词: 车流控制; 数据驱动; 车路协同; 深度强化学习; 网联自动车; 混合交通流

中图分类号: U491.5+4

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0700

引用格式: 张辰, 徐云雯, 李德伟. 车路协同环境下数据驱动的混合交通流速度调控方法[J]. 控制与决策, 2024, 39(9): 2950-2958.

Data-driven speed control method for mixed traffic flow in vehicle-road cooperative environment

ZHANG Chen, XU Yun-wen[†], LI De-wei

(Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Aiming at the problem of mixed traffic flow control in the context of development of cooperative vehicle infrastructure systems and autonomous driving technology, a speed control method of controlled connected autonomous vehicle based on road side unit (RSUs) to regulate the speed of macroscopic traffic flow is proposed. Firstly, this paper describes the adjustment process of a single controlled connected autonomous vehicle to the traffic state as a Markov decision process. A dynamic grid is designed to measure the influence of the controlled vehicle on the overall traffic flow. Considering the influence of the controlled vehicle on the traffic flow in the dynamic grid, the response speed to the expected target and the safety factor of the overall traffic flow, the reward function is constructed. The deep deterministic policy gradient (DDPG) algorithm is adopted for policy optimization. Considering the diversity of the traffic environment around the controlled connected autonomous vehicle, a control strategy model cluster based on key parameters is proposed for RSUs to select in real time according to traffic flow states. Secondly, for the real-time adjustment of traffic average speed, the division method of vehicle clusters in road sections and the timing handover strategy of control rights for multiple roadside units to control the particular vehicle cluster are proposed. The RSUs track vehicles within the signal coverage area in a vehicle cluster, and calculate the expected control signal in real time, then send it to the controlled vehicles. Finally, the proposed method is verified in different scenarios, which can smoothly and efficiently regulate the speed of traffic in multiple scenarios.

Keywords: traffic flow control; data-driven; cooperative vehicle infrastructure system; deep reinforcement learning; connected autonomous vehicle; mixed traffic flow

0 引言

随着车联网和自动驾驶等技术的发展, 车路协同的应用场景不断丰富, 车流调控手段已经不局限于传

统的交叉口信号灯控制等方法^[1-2], 利用车路协同技术从微观车辆层面协调车流平均速度成为可能, 有助于在车联网环境下提升路网通行能力.

收稿日期: 2023-05-22; 录用日期: 2023-09-14.

基金项目: 国家重点研发计划项目(2020YFB1600400); 国家自然科学基金项目(62003210, 62333015).

责任编辑: 郭戈.

[†]通讯作者. E-mail: willing419@sjtu.edu.cn.

传统快速路交通流动态控制方式包含拥堵吸收驾驶策略和可变限速控制^[3]。目前,拥堵吸收控制往往应用在单车道场景中;可变限速控制在车辆行驶过程中增添了速度控制节点^[4],但可变限速方式无法区分微观车辆间的差异,只能对到达特定位置处的车辆进行约束,且当驾驶者从自身通行效率最大化的目标出发,往往以最大限速行驶,会减弱限速的控制效果^[5]。

网联自动驾驶车辆的出现增加了车流调控自由度,通过调整车辆速度影响周边车流,进而控制整体车流平均速度,是未来交通控制的有效手段^[6-7]。目前,网联车群体控制的研究集中在多个车辆间的协同控制,如研究特定场景下(路口、匝道等)由自动驾驶车辆所形成的车队(platoon)控制。大部分工作以车队任务为导向来设计控制目标^[8]。多种控制方法在车辆的动力学控制、编队构成以及车队的分布式控制中得到应用,如非线性控制^[9]、在随机网络中基于观测器设计控制率^[10]、随机动态规划、协同控制^[11]等。车队控制是基于部分车辆利益层面的优化问题,在此过程中很少讨论如何改变交通流的整体状态。

随着车路协同技术的发展,实时环境感知可由路侧设备(road side unit, RSU)完成,并将实时感知数据处理后迅速发布给其覆盖范围内的行驶车辆。此外,RSU可以根据全局路网或者小区域内的车辆拥堵分布,为自动驾驶车辆规划控制方案并实时下发。因此,当路侧设备能够接管其信号覆盖范围内所有受控自动驾驶车辆的控制时,路面的车流便处于可控的状态^[12]。如何调控混合交通流中的受控车辆使整体车流达到期望状态,是实现微观车辆调节混合车流的关键。

在车辆运动过程中,由于人驾驶车辆的随机性,受控车辆对周边车辆影响的控制模型难以建立。车辆加速受到前方车辆的约束,减速对整个车流的影响与当前车流的平均车距、平均车速以及该车辆于路段的相对位置等因素密切相关^[13],难从机理上建立准确的控制模型。在现有文献中很多学者利用数据驱动方式进行控制,其在面临输入输出维数较高以及难以建模等自动驾驶问题时有较好的表现。如车辆的跟踪控制,设计自动驾驶车辆换道策略^[14]等。强化学习算法通过智能体与其所处环境的信息交互,以马尔可夫决策过程为基础通过不断试错探索出当前环境下的最优决策^[15]。根据动作空间的特点,有离散动作空间和连续动作空间两类强化学习算法。确定性策略梯度(deterministic policy gradient, DPG)

和深度确定性策略梯度(deep deterministic policy gradient, DDPG)是具有代表性的连续动作空间强化学习算法,它们通过训练值函数网络来估计当前环境下采取特定动作获得的 Q 值,将策略网络与值函数网络相连,在策略网络进行优化时,评价网络便可提供优化的梯度。DDPG用策略网络弥补了离散型强化学习方法的不足,基于确定性的策略梯度给出最优策略。

基于上述考虑,本文针对快速路车流系统,研究通过控制车流中的受控车辆调整车流平均速度的问题。基于马尔科夫决策过程,对单车变速对周边车流的影响进行建模,利用强化学习DDPG方法构建相应的评价和策略网络,实现交通流的速度调节;基于路侧RSU实现对移动车流的跟踪控制,RSU以簇为单位对信号覆盖范围内的车辆进行标记与跟踪,通过路侧控制模块实时计算受控车辆的期望控制信号并下发,多个RSU间以协同传递的方式实现对车辆簇的控制。

1 问题描述

面向由自动驾驶网联车和人工驾驶车组成的混合车流,本文提出利用路侧感知单元RSU对自动驾驶网联车的运动状态进行实时控制,通过车辆之间运动状态的耦合与相互约束,达到调节整体车流状态的目标。下面对研究场景进行描述。

本文以城市某快速路段为研究对象,假设路面车辆类型、位置、速度等信息均可被路侧RSU实时感知,RSU之间通过相互通信实现对路段的全息感知^[16]。路侧RSU可以与进入其通信范围内的自动驾驶网联车进行通信。记路段上路侧RSU数量为 n ,则路段RSU集合记为 $U = \{U_1, U_2, \dots, U_n\}$ 。在 k 时刻,进入路侧RSU U_i 通信范围内的所有车辆集合记为 $V_{(i,k)} = \{C_{(i,k)}, H_{(i,k)}\}$,其中 $C_{(i,k)}$ 为受控自动驾驶网联车(controlled connected autonomous vehicle, CCAV)集合, $H_{(i,k)}$ 为路面人驾驶的非受控车辆HV集合。当路段接收到车流调控指令时,路侧RSU U_i 向 C_i 集合内的受控车辆发送控制信号,实现对 V_i 的控制。如图1所示,路侧RSU通过信息交互对路面车流进行全维感知。当路段接收到车流调控指令时,路侧RSU将会发送控制信号给通信范围内受控自动驾驶网联车,使其改变当前的运动状态。同时,周边非受控车辆HV都会受到相应影响,进而改变整体车流。因此,本文研究目标是在路侧RSU协同通信的背景下,通过控制局部CCA V的速度,将路段混合车流的速度调节到期望平均速度 v^* 。

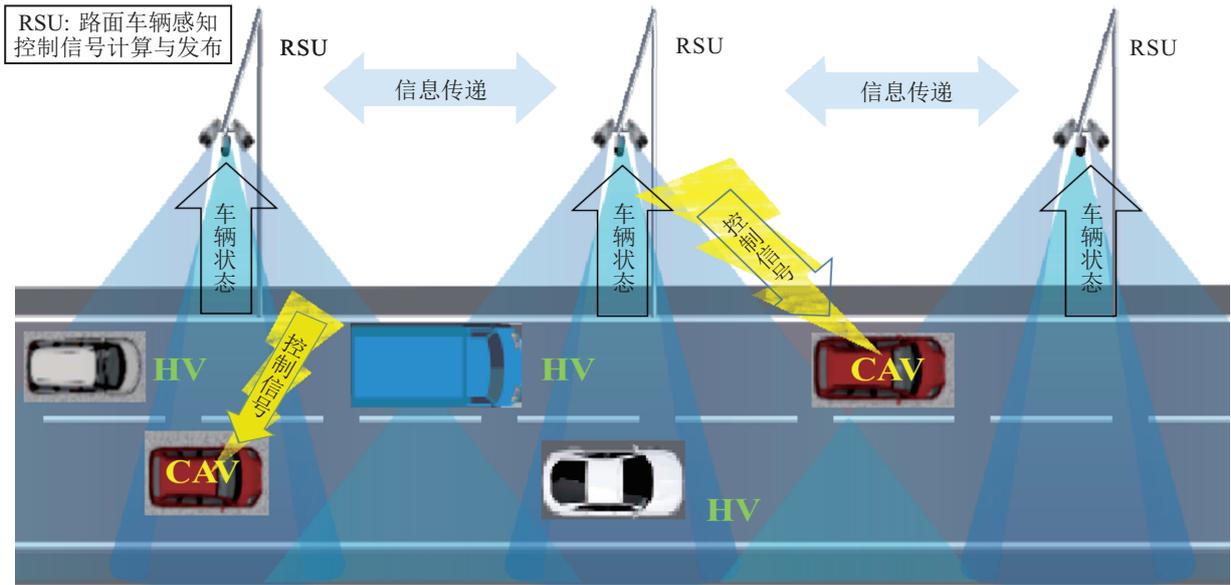


图1 车流调控场景

2 面向混合车流调控的CCAV状态控制方法

由于混合车流中不同位置、不同速度车辆的速度变化对于整体车流的影响作用具有不确定性,难以从机理上建立精确的数学模型进行描述.本节将CCAV对车流进行调节的问题描述为马尔科夫决策过程(Markov decision process, MDP),其中每个CCAV分别作为一个独立的agent.通过改变单个agent的行驶速度,观察其与路面其他车辆的交互作用,得到周边交通车流状态的变化规律,进而学习每个agent在混合车流速度调控目标驱动下的行驶策略.下面首先对单个agent的状态、动作以及报酬函数的定义进行详细描述.

2.1 状态

对于每一个CCAV,其速度的改变会导致周边车辆的速度、加速度也随之动态变化.从调节整体混合车流的控制目标出发,本文将每个CCAV的状态定义为被其影响的周边车流的状态.首先对单个CCAV减速能够影响的车流范围进行计算.设受控自动网联车 C_m 在第 k 个周期的行驶速度为 v_k^m ,该车在第 k 个周期(对应时刻为 t_k)处于 U_i 感知范围内,且 U_i 感知范围内车流平均速度为 $v_{k,i}$.在车辆移动方向, C_m 的影响范围与其造成的减速波移动的速度有关.当 C_m 在速度相对稳定的车流中(即理想情况下所有车辆均以同一速度行驶或车速方差很小)进行减速时,其上游和周边车辆会跟随减速,在车队中形成减速波,并向后传播.考虑路段车速在短时间内不发生剧烈变化的特点,根据不同波之间的跃变条件,近似计

算以行进车流为参考系的相对波速为

$$u_k^m \approx \bar{v}_{k,i}^m - \bar{v}_{k,i}, \quad (1)$$

其中 $\bar{v}_{k,i}^m$ 代表以 C_m 相邻两个控制周期的平均速度,利用 C_m 的平均速度与 U_i 感知范围内车流平均速度之差近似代替波在车流中的传播速度.由这一近似公式可知,当 C_m 速度未发生改变时,将跟随车流的平均速度,此时波速等于0,符合实际情况.

规定网格长度的最小值为包含受控车前后各一辆车的车长度 L_0 ,并将 C_m 移动方向上的影响范围记为 L_m ,则

$$L_m = \max\{|u_k^m| \cdot T, L_0\}, \quad (2)$$

其中 T 为 U_i 的感知信号的采样周期.如图2所示,虚线方框是以 L_m 为长度,3车道为宽度的动态网格 N_m ,即 N_m 包含 C_m 影响范围内的车辆集合.在本文讨论的双向六车道场景中,当 C_m 处于中间车道时,两侧车辆均在影响范围内,此时影响范围横跨三车道;当 C_m 处于两边车道时,影响范围包含自身车道与中间车道.

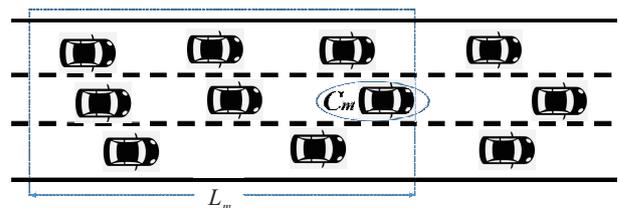


图2 动态网格

基于式(2)定义的网格长度,将CCAV减速后的状态定义为

$$S_k^m = (v_k^m, \bar{v}_k^m, a_k^m, \bar{a}_k^m, s_{k,v}^m)^T, \quad (3)$$

其中路侧RSU U_i 根据动态网格 N_m 大小, 计算以受控车辆 C_m 为起始的网格内所有车辆的平均速度和加速度, 分别记为 \bar{v}_k^m 、 a_k^m 、 \bar{a}_k^m , 网格内速度方差为 $s_{k,v}^m$.

2.2 动作

对于任一个受控CCAV C_m , 设在时刻 k 其期望速度为 $v_{k,i}^m$, 车流期望速度为 v^* , 则其可行动作空间定义为

$$A_k^m = \{v_{k,i}^m | v^* - \sigma \leq v_{k,i}^m \leq v_{\max}\}. \quad (4)$$

其中: 实时期望速度是满足上述范围的连续变量, σ 为阈值, 在减速场景中, 车辆期望速度的最大值 v_{\max} 为该车所处网格中的初始平均速度 \bar{v}_0^m . 在时刻 k , 路侧RSU U_i 对进入信号范围内的车辆 C_m 发送参考速度信号 $v_{k,i}^m$. 接到信号后, C_m 自动驾驶系统根据当前路况, 在安全条件下调整车辆速度达到期望目标 $\bar{v}_{k,i}^m$, U_i 下发动作周期间隔与RSU感知信号采样周期相同, 均为 T .

2.3 报酬函数

报酬函数主要从个体CCAV对周边车流控制效果的角度进行设计, 具体包含对网格车流的影响程度、对期望速度的响应快慢以及整体车流的安全性3个方面. 下面分别对这3方面进行详细阐述.

对周边车流的影响程度通过网格 N_m 内的平均速度 \bar{v}_k^m 与车流期望速度 v^* 的绝对值差来衡量, 由于报酬函数为极大化函数, 所以将具体报酬表示为

$$J_k^1 = -|\bar{v}_k^m - v^*|. \quad (5)$$

对期望目标的响应速度的调整通过网格 N_m 内车辆平均加速度指标来体现. 本文采用自适应的思路进行设计. 报酬项具体形式如下:

$$J_k^2 = \begin{cases} |\bar{a}_k^m|, & |\bar{v}(k) - v^*| - \epsilon \geq 0; \\ (\bar{a}_k^m + a_0)^{-1}, & |\bar{v}(k) - v^*| - \epsilon < 0. \end{cases} \quad (6)$$

其中 a_0 是为了保证 J_k^2 连续所取的正常数. 当 N_m 内的平均速度与期望值之差较大时, 即满足 $|\bar{v}(k) - v^*| - \epsilon \geq 0$, 需增大整体加速度使其向目标状态过渡, 此时 $J_k^2 = |\bar{a}_k^m|$, 加速度越大, 所获得的报酬越大; 当车辆接近期望平均速度时, 即 $|\bar{v}(k) - v^*| - \epsilon < 0$ 时, 需减小平均加速度, 防止状态在接近目标 v^* 时出现震荡, 此时 $J_k^2 = (\bar{a}_k^m + a_0)^{-1}$, 加速度越小, 所获得报酬越大.

考虑到调速过程车辆的驾乘舒适性和状态变化的平稳性, 构造如下惩罚项避免车辆陡然减速引起的安全问题:

$$J_k^3 = J_k^1 \cdot \text{Relu}(\epsilon_v - |v_k^m - v_{k-1}^m|). \quad (7)$$

其中: ϵ_v 为阈值, $\text{Relu}(x) = x + |x|$, $\epsilon_v - |v_k^m - v_{k-1}^m|$ 衡量了 C_m 前后两次控制时刻速度波动幅度, 控制过程要求相邻两次控制周期车辆速度变化量在阈值范围内. 若车辆速度陡降或陡增, 则会造成车流状态震荡, 存在安全隐患, 定义这样场景的报酬为0. 当 C_m 的车速变化在阈值之内时, N_m 内的平均速度越接近目标车速, C_m 的车速变化越小, 所得报酬 J_k^3 越大.

综上, 在 k 时刻车辆 C_m 的总报酬函数为

$$R_k^m = \alpha_1 J_k^1 + \alpha_2 J_k^2 + \alpha_3 J_k^3, \quad (8)$$

其中 α_1 、 α_2 、 α_3 为正的权重参数.

2.4 CCAV梯度策略参数更新算法

在受控车辆的速度控制中, agent利用深度确定性梯度策略算法(DDPG)进行控制策略更新. 该算法基于确定性的策略梯度和Actor-Critic框架, 可以用于连续控制量的动作空间, 其中Actor模块中输入为 C_m 在网格 N_m 内的状态值, 输出为 C_m 的实时期望速度; Critic模块输入为状态 S_k^m 和期望速度 v_i^m , 输出为报酬值.

用 $\pi_m(s, a)$ 表示 C_m 的行为策略, 即在以其自身构造的移动网格的状态 $S_m = s$ 下采取的控制期望速度 $A^m = v_i^m = a$ 的概率分布, 即 $\pi_m(s, a) = P[A^m = a | S^m = s]$. 用值函数 $Q_m^\pi(s, a)$ 表示 C_m 在网格状态为 s 的情况下根据策略 π_m 采取的控制量 a , 在累计折扣 γ 下报酬的期望. 其中 Q_m^π 定义为

$$Q_m^\pi(s, a) = E_\pi[R_{k+1}^m + \gamma Q_m^\pi(S_{k+1}^m, A_{k+1}^m) | S_k^m = s, A_k^m = a]. \quad (9)$$

在迭代过程中, CCAV通过获得最大期望报酬的优化目标更新策略 μ_m^* , 通过策略 μ_m^* 得到状态 S_k^m 下的 C_m 期望车速 A_k^m . 从估计的目标和目标网络学习更新参数以保持其估计目标的稳定, 防止目标发散.

3 路段混合车流速度调控方法

为了实现车流平均速度整体控制, 需要路段上各RSU的协同交互. 本文首先将路面车流划分为若干车辆簇VC(vehicle cluster). 路侧RSU跟踪不同VC中的CCAV, 从控制策略模型簇中选择相应控制模型, 下发控制信号以达到整体控制目标.

3.1 车辆簇划分

车辆簇内的车辆通常具有位置分布紧密, 实时速度相近的特点. 以车辆簇为整体进行调控可以提

高车流速度控制效率,便于车辆跟踪控制.为了保证调控的有效性,本文对车辆簇内 CCAV 的渗透率和车辆簇长度分别设定阈值 p_0 和最大长度 L ,以保证含有 CCAV 的车辆簇中,渗透率和簇长度满足给定阈值.本文应用 size constrained clustering 算法^[17]进行聚类,每辆车的特征表达为 (l_{no}, p_r, v) . 其中: l_{no} 为车辆所处车道编号, p_r 为车辆在车道中的位置, v 为车辆速度. 设路段上所有车辆被划分为 N 个车辆簇,按顺序分别记为 VC_1, VC_2, \dots, VC_N ,其中每辆车只属于一个车辆簇.

随着车流在路面移动,每个路侧 RSU 以车辆簇为单位对其中的 CCAV 进行跟踪控制,通过交互得到对应车辆簇中 CCAV 的信息.

3.2 基于关键参数的RSU控制策略模型簇

在第2节讨论的控制方法中,针对单辆 CCAV 设计了控制器. RSU 在调节车辆簇速度的过程中,不同车辆簇中 CCAV 的渗透率不同,以及不同的控制目标都需要选择相应的控制模型,因此基于不同场景和需求构建控制策略模型簇供 RSU 实时选择.模型簇包含 CCAV 渗透率、控制目标速度、平均车头间距这3个关键参数.

随着渗透率的不断增大,CCAV 的控制能力随之增强;反之,当 CCAV 渗透率较低时,其他车辆可以换道加速驶离,使得控制失败.将 CCAV 渗透率 p 划分为多个范围 $[p_0, p_1), [p_1, p_2), \dots, [p_{X-1}, p_X]$.

报酬函数优化目标包含了实时网格速度与控制目标的差值 J_k^1 . 在不同的控制目标 v^* 下,相同大小的 J_k^1 会对应不同的期望车速 A_k^m . 因此,需将控制目标 v^* 划分为多个区间 $[v_0^*, v_1^*), [v_1^*, v_2^*), \dots, [v_{Y-1}^*, v_Y^*]$ 进行考虑.

平均车头间距也是影响模型泛化能力的重要因素.在周边环境状态 S_k^m 相同的情况下,在车头间距越大,动作序列的变化幅度应越大,减少车流簇中的换道行为,使得整体车流能够快速收敛到目标值.因此,将不同的车头间距 h 划分为多个场景 $[h_0, h_1), [h_1, h_2), \dots, [h_{Z-1}, h_Z]$. 在上述3个维度下,生成模型簇 $M_{(p,v^*,h)}$.

3.3 RSU信息传递与计算流程

本节讨论 CCAV 在不同 RSU 间的控制权交接问题. 设 $k = 0$ 时刻,路侧 RSU 接收到控制指令将车辆簇 VC_{p^*} 由平均速度 v_0 下调到 v^* . 假设此时车辆簇 VC_{p^*} 处于 RSU U_i 的通信范围之内,初始车辆簇集合

为 $VC_{p^*} = \{C_{p^*}, H_{p^*}\}$. U_i 通过感知设备获取 C_{p^*} 内所有车辆的当前状态 $\{S_{0,p}^1, S_{0,p}^2, \dots, S_{0,p}^M\}$,根据当前车辆簇中的 CCAV 渗透率、目标速度和平均车头间距,从控制策略模型簇 $M_{(p,v^*,h)}$ 中选取对应的策略网络 $\mu_1, \mu_2, \dots, \mu_M$,向集合 C_{p^*} 中的车辆下发初始控制信号 $v_0^1, v_0^2, \dots, v_0^M$.

随着车辆的移动,车辆簇 VC_{p^*} 会离开 U_i 的通信范围,进入 U_{i+1} 所在区域. 设 U_i 与 U_{i+1} 之间覆盖范围的重叠区域为 O_i ,该区域为控制权的交接区. 将 U_i 的信号覆盖范围分为3个部分:与 U_{i-1} 的信号重叠区 O_{i-1} ,信号不重叠区 R_i ,与 U_{i+1} 的信号重叠区 O_i . 控制流程见图3.

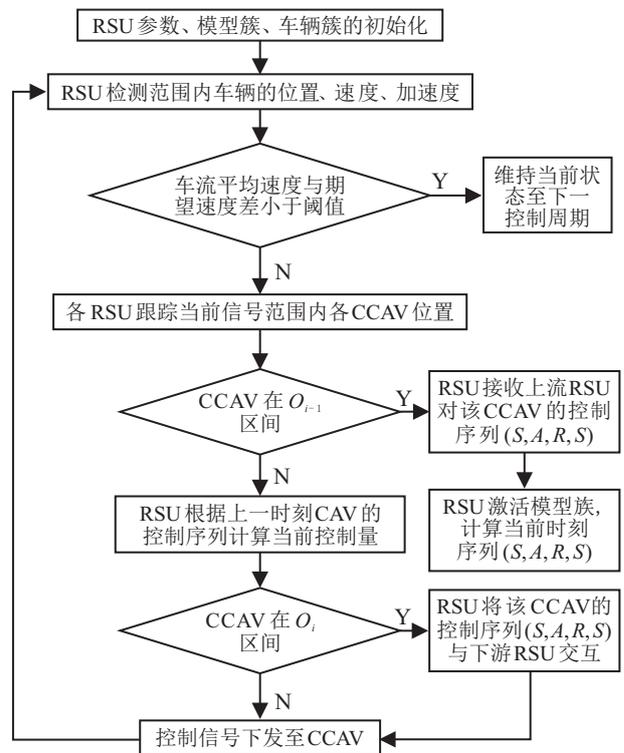


图3 RSU控制流程

在 k 时刻,当车辆簇 C_{p^*} 内车辆 C_m 在 O_{i-1} 的信号范围内,车辆 C_m 的控制信号 v_k^m 由上游的 U_{i-1} 计算得出, U_{i-1} 将对应的序列 $\{(S_{k-1}^m, A_{k-1}^m, R_{k-1}^m, S_k^m), (S_{k-1}^m, A_{k-1}^m, R_{k-1}^m, S_k^m), \dots\}$ 传送给 U_i ,直到 C_m 移动到 R_i 范围内,控制权交接给 U_i , $\{v_k^m, v_{k+1}^m, \dots\}$ 由 U_i 计算得出. 随着 C_m 移动至 O_i 范围内, U_i 在计算控制信号的同时,将 C_m 的信息 $\{(S_k^m, A_k^m, R_k^m, S_{k+1}^m), (S_{k-1}^m, A_{k-1}^m, R_{k-1}^m, S_k^m), \dots\}$ 传送给 U_{i+1} ,准备进行下一轮控制权的交接.

3.4 整体控制流程

系统的控制框架如图4所示,图4左侧展示了路段车辆状况,当路段 RSU 接受到车流调控信号后,在

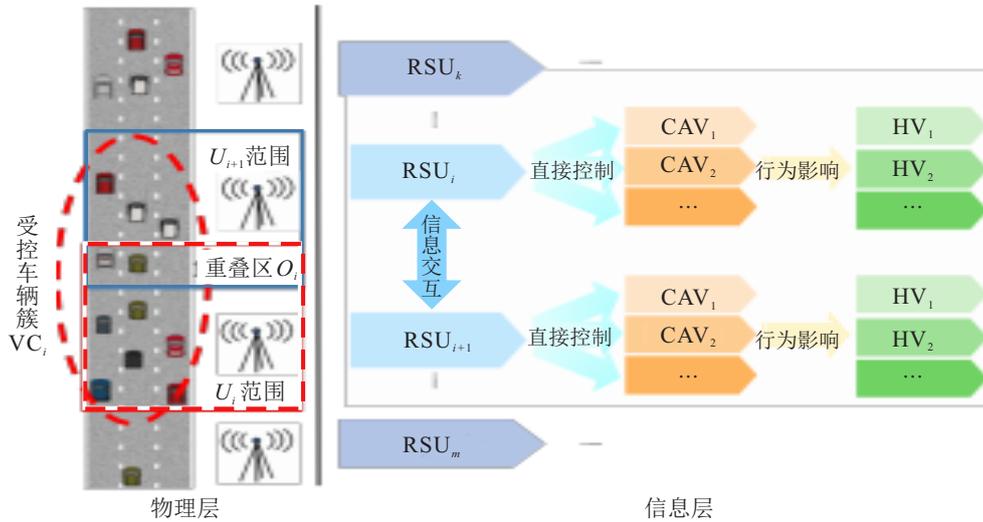


图4 系统控制框架

信息层将车流根据上文车辆簇划分方法划分为多个车辆簇 VC_1, VC_2, \dots, VC_n , 分簇进行控制(图4中红色虚线内为示例的车辆簇 VC_i).

假设目前 VC_i 中有 M 辆车进入 U_i 信号范围(图4中红色方框内), 其中 N 辆 CCAV 根据对应的 CCAV 渗透率, U_i 激活模型簇 $M_{(p,v^*,h)}$, 利用实时采集到的状态计算各自对应的动作值, 并由 U_i 向这 N 辆 CCAV 分别发送控制信号(每辆车的实时期望速度). 该段车流中处于不同位置的 CCAV 接到控制信号后, 自身的自动驾驶系统调整动力向目标运动状态进行改变, 同时由于受控车流的车辆间存在位置、速度和加速度间的约束, 范围内 HV 也会因此改变自身的运动状态. 在车流进入下一个控制范围(图4中蓝色框内)时, U_i 和 U_{i+1} 进行控制信号和感知信息的交互, 在重叠区 O_i 交接控制权, 多个 RSU 形成对路面车流的整体控制, 使其达到目标状态.

4 实验验证与分析

本文基于 SUMO/CACC 仿真平台搭建3车道快速道路环境以验证算法的效果. 实验路段前2km为缓冲区, 后侧4km为控制测试区. 分别应用本文提出的基于 DDPG 策略更新方法、DQN 方法和 CCAV 自适应控制方法在不同交通环境(车辆簇不同规模、不同密度、不同渗透率)中进行车流调速过程的仿真分析.

仿真环境中, HV 采用 IDM 模型模拟人类驾驶行为. CCAV 采用固定车头时距模式, 接收 RSU 发送的速度信息.

本文根据模型簇关键参数, 选取处在车流不同位置的受控车作为强化学习算法的训练对象,

得到多个数据模型. 强化学习采用 DDPG 和 DQN 两种方法进行比较. 其中: DDPG 动作空间阈值为 4km/h, 每轮最大训练时间为 300 s, 采样周期为 1 s; 为方便测试, 将 DQN 动作空间离散化, 动作空间为 30 km/h ~ 60 km/h, 速度间隔为 3 km/h, 共 11 个动作状态. 具体参数如表1所示.

表1 强化学习算法参数

算法参数	DDPG	DQN
采样时间/s	1	1
每轮训练时间/s	300	300
Actor 网络学习率	0.0001	无
Critic 网络学习率	0.0002	0.0002
更新参数 τ	0.001	0.001
采样参数 N	128	128
最大训练轮数	5000	5000

模型簇中目标速度 v^* 范围是 30 km/h ~ 65 km/h, 每隔 5 km/h 划分一个区间; 渗透率 p 范围是 20% ~ 80%, 每隔 10% 划分一个区间; 车头间距范围 h 是 5 m ~ 35 m, 每隔 5 m 划分一个区间. 因此, 模型簇中包含模型数量为 $|p| \times |v^*| \times |h| = 8 \times 7 \times 7 = 392$.

实验同时验证两个车辆簇 VC_1 和 VC_2 . VC_1 内车辆速度均值服从高斯分布, 其均值为 60 km/h, 方差为 5; 车辆簇目标平均速度为 $v_1^* = 45$ km/h. VC_2 车流簇中车辆速度均值为 50 km/h, 方差为 2; 目标平均速度 $v_2^* = 35$ km/h. 初始平均车头间距分为 15 m、25 m 和 35 m 三种情况, 车流在缓冲区分别达到上述稳定状态后进入控制测试区. 车辆簇的规模分为 30 辆车、40 辆车、50 辆车的 3 种场景. 在不同 CAV 渗透率的情况下由 RSU 实时选取模型, 进行多轮测试, 统计所有车辆的平均速度关于控制步数的变化曲线.

首先比较3种不同的车辆状态控制模型的控制效果. 其中在应用DDPG和DQN这两种数据驱动方法的实验中,每个控制周期,RSU会根据受控车流的状态利用不用的数据驱动模型计算控制量,并实时向CCAV发送对应的参考速度信号. 在自适应控制算法中,RSU会直接将设定的车流目标平均速度 v^* 发送给每辆CCAV,CCAV的自动驾驶系统会调整自身的速度和加速度将其速度状态直接向 v^* 调整. 图5 (CCAV渗透率40%,平均车头间距15 m)和表2展示了规模为40辆车的车辆簇 VC_1 在给定场景下多次实验3种不同方法的平均效果.

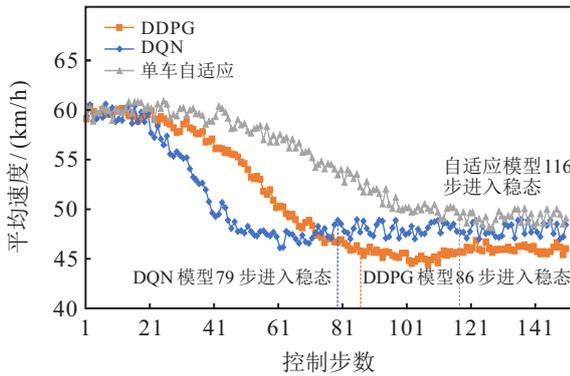


图5 不同算法的平均速度变化曲线

表2 50辆车、间距25 m、CAV渗透率35%的仿真结果

评价指标	DDPG	DQN	自适应
到达稳态所需控制步数	215	302	无
稳态平均速度/(km/h)	45.9	48.2	无

由图5和表2可以发现,基于强化学习的两种CCAV状态控制模型均能成功改变车流的平均速度. 在40辆车的规模,CCAV渗透率40%的场景下,基于DDPG的车辆控制模型达到了设定的目标值. DQN的动作增大会大大增加搜索空间,降低算法效率和效果,使得控制信号的计算时间大于控制周期;由于动作空间受限,控制效果容易发生震荡,平均速度跳变明显. 因为强化学习的算法在训练阶段存储了大量的经验,所以针对40辆车和50辆车的车辆簇仍能完成调速任务. 单车自适应控制方法在40辆车的情况还可以使车流进入稳态,而车辆数再度增大便难以胜任,这也体现出利用数据构建模型对于解决这个问题的优势.

如图6所示,实验场景在30辆车流规模,初始平均车距为15 m的情况下,随着受控网联车的渗透率不断提高,车流达到新稳态的时间减小. 当渗透率减小时,稳态速度均值与目标差距增大,调控效果降低.

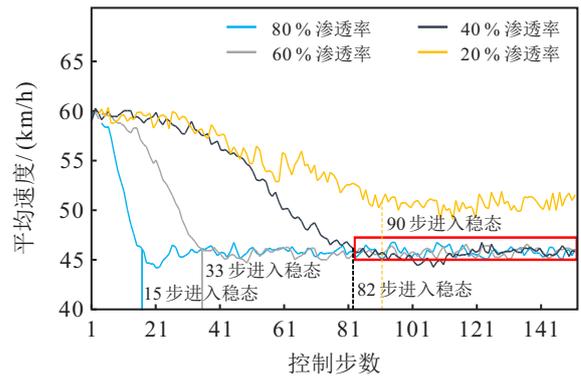
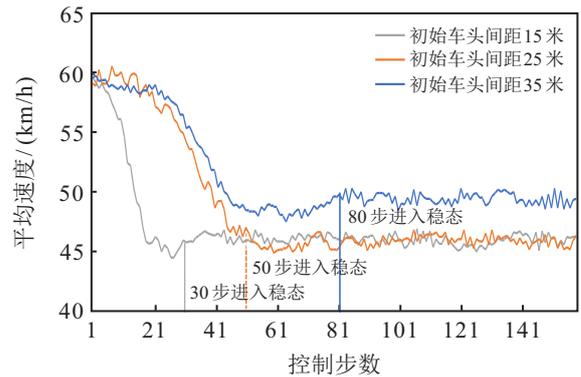
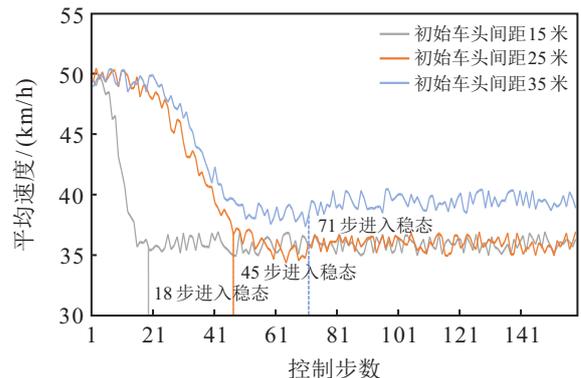


图6 不同CCAV渗透率场景平均速度变化曲线

如图7所示,进入稳态的时间随车辆的初始平均车距增大而增大. 当初始平均车距为35 m时,最终受控车辆簇平均速度增高,效果下降. 原因是当路面车距大,车流分布较散时,前方车辆减速,后侧非受控车变道超车机会较大,当非受控车超车后,失去了前方约束,其速度无法受控;而当车距小于一定距离时,换道频率降低,平均速度也随之下降.



(a) 30辆车 VC_1 , CCAV渗透率45%



(b) 30辆车 VC_2 , CCAV渗透率45%

图7 不同初始平均车距场景平均速度变化曲线

在初始平均车距25 m和CCAV渗透率45%的场景下研究车流规模的影响. 如图8所示,在相同初始平均车距、同样CCAV渗透率的车辆簇 VC_1 调节场景中,达到新稳态的耗时随着车辆簇的规模增大而增加. 原因是车辆簇规模增大,减速状态传递所花时间

也随之增大. 在高渗透率下, 即使有车距增大的影响, 由于受控车辆数较多, 非受控车辆周边的约束较大, 算法的效果依然保持, 保证设定目标速度可以达到.

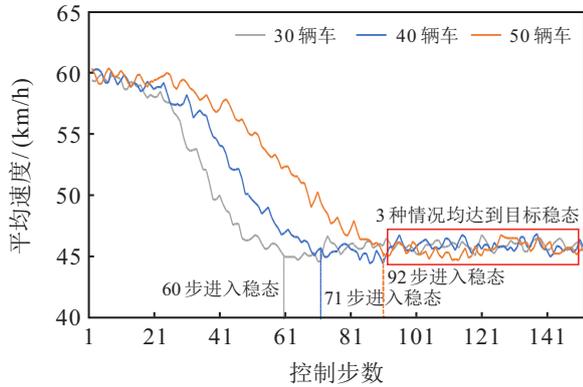
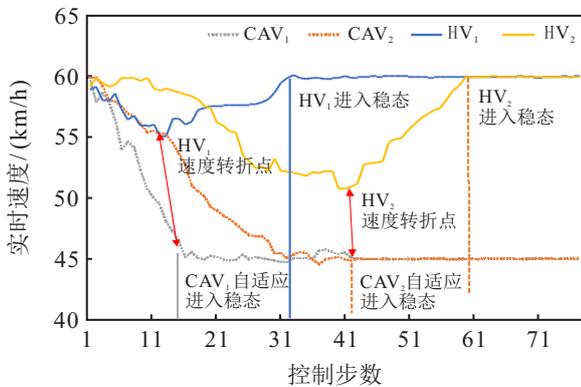
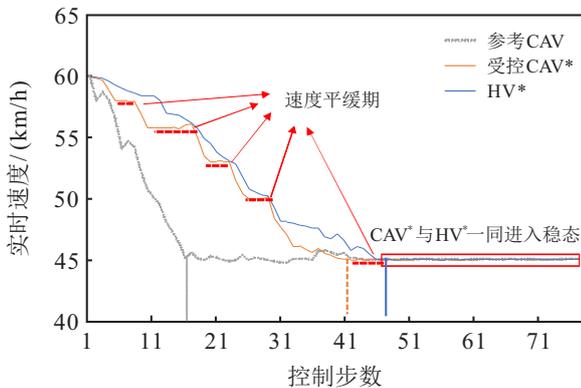


图8 不同受控车流规模平均车距场景平均速度变化曲线

在30辆车流规模、初始平均车头间距25米、CAV渗透率45%的情况下研究不同车辆的速度变化, 如图9所示. 图9(a)给出减速车辆CAV₁、CAV₂以及各自后方HV₁、HV₂的行驶速度变化曲线, 是CAV₁与CAV₂的期望速度被直接设置为45 km/h, 由车内自动驾驶系统根据当前环境自适应降低车速的控制场景; 图9(b)是CAV根据路侧RSU控制器实时给出的参考速度信号进行驾驶的情况下, 其后方车辆的速度变化曲线. 图9(a)中, CAV₁由于目标被直接设置



(a) 减速车辆CAV及其后方HV的行驶速度变化曲线



(b) CAV后方车辆的速度变化曲线

图9 不同车辆速度变化曲线

为45 km/h, 车辆短时间内下降到设定目标, 而其后方车辆HV₁为了防止追尾在前11控制步内也随之降低车速. 当CAV₁达到新稳态平均速度后, HV₁提升了速度实现超车, 降低控制效果. 同样地, CAV₂仅通过自身自动驾驶系统降低速度达到设定值, 其后方车辆在HV₂达到稳态前跟随减速, 之后回归初始速度, 也导致了未达成整体调速目标. 图9(b)中, 受控CAV*所对应的曲线速度下降过程呈阶梯式, 速度每下降一段时间后, 其速度的改变量会减小, 进入速度平缓期(即虚线阶段); 当多辆CAV均通过此方式降速, 且各自速度平缓期错开, 可以减少周边HV的超车机会, 使得HV跟随CAV减速, 实现平均速度下降.

5 结论

本文实现了车路协同系统下对道路车流的速度状态控制. 通过路侧单元RSU实时感知路面车辆信息, 路侧采用基于数据驱动的方法对改变CAV实时期望速度、调节整体车流的运动状态的场景进行建模, 得到了如下结论: 车辆簇初始平均车头间距越小, 车辆簇CAV渗透率越高, 算法的控制效果越好. 将不同控制方法在不同场景中验证, 结果表明: 相比于DQN和无模型的自适应控制, 基于DDPG的强化学习控制效率和控制精度都相对较高. 本文提出的方法在不同的车辆数、不同渗透率情况下均具有良好的适应性.

本文提出的方法从全局角度高效控制交通流, 通过以控制受控CAV的方式调节整体交通流的平均速度, 对未来城市交通流的精细化控制有参考意义.

参考文献(References)

[1] 左志强, 刘正璇, 王一晶. 基于车路云一体化的混合交通系统优化控制综述[J]. 控制与决策, 2023, 38(3): 577-594.
(Zuo Z Q, Liu Z X, Wang Y J. A survey of optimal control for mixed traffic system with vehicle-road-cloud integration[J]. Control and Decision, 2023, 38(3): 577-594.)

[2] Guo G, Zhang T Q. A residual spatio-temporal architecture for travel demand forecasting[J]. Transportation Research Part C: Emerging Technologies, 2020, 115: 102639.

[3] Lu Q, Tettamanti T. Impacts of connected and automated vehicles on freeway with increased speed limit[J]. International Journal of Simulation Modelling, 2021, 20(3): 453-464.

[4] Du S M, Razavi S. Variable speed limit for freeway work zone with capacity drop using discrete-time sliding mode

- control[J]. *Journal of Computing in Civil Engineering*, 2019, 33(2): 04019001.
- [5] Li D, Fooladivanda D, Martínez S. Data-driven variable speed limit design for highways via distributionally robust optimization[C]. 2019 18th European Control Conference. Naples, 2019: 1055-1061.
- [6] Moradi H, Sasaninejad S, Wittevrongel S, et al. Proposal of an integrated platoon-based Round-Robin algorithm with priorities for intersections with mixed traffic flows[J]. *IET Intelligent Transport Systems*, 2021, 15(9): 1106-1118.
- [7] Zhang T Q, Guo G. Graph attention LSTM: A spatiotemporal approach for traffic flow forecasting[J]. *IEEE Intelligent Transportation Systems Magazine*, 2022, 14(2): 190-196.
- [8] Wang P W, Jiang Y L, Xiao L, et al. A joint control model for connected vehicle platoon and arterial signal coordination[J]. *Journal of Intelligent Transportation Systems*, 2020, 24(1): 81-92.
- [9] Yan M D, Ma W R, Zuo L, et al. Distributed model predictive control for platooning of heterogeneous vehicles with multiple constraints and communication delays[J]. *Journal of Advanced Transportation*, 2020, 2020: 1-16.
- [10] Liang X L, Zhang Y X, Yang G T. Platoon control design for unmanned surface vehicles subject to input delay[J]. *Scientific Reports*, 2021, 11(1): 1-12.
- [11] Wen S X, Guo G. Observer-based control of vehicle platoons with random network access[J]. *Robotics and Autonomous Systems*, 2019, 115: 28-39.
- [12] Gong S Y, Du L L. Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles[J]. *Transportation Research Part B: Methodological*, 2018, 116: 25-61.
- [13] Wang J W, Zheng Y, Xu Q, et al. Controllability analysis and optimal control of mixed traffic flow with human-driven and autonomous vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(12): 7445-7459.
- [14] Hoel C J, Wolff K, Laine L. Automated speed and lane change decision making using deep reinforcement learning[C]. 2018 21st International Conference on Intelligent Transportation Systems. Maui, 2018: 2148-2155.
- [15] 温广辉, 杨涛, 周佳玲, 等. 强化学习与自适应动态规划: 从基础理论到多智能体系统中的应用进展综述[J]. *控制与决策*, 2023, 38(5): 1200-1230.
(Wen G H, Yang T, Zhou J L, et al. Reinforcement learning and adaptive/approximate dynamic programming: A survey from theory to applications in multi-agent systems[J]. *Control and Decision*, 2023, 38(5): 1200-1230.)
- [16] 龙学军, 谭志国, 高枫. 多传感器融合路侧感知技术应用现状分析[J]. *中国交通信息化*, 2021(10): 137-140.
(Long X J, Tan Z G, Gao F. Analysis of application status of multi-sensor fusion roadside sensing technology[J]. *China ITS Journal*, 2021(10): 137-140.)
- [17] Li X L, Zhang H, Wang R, et al. Multiview clustering: A scalable and parameter-free bipartite graph fusion method[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(1): 330-344.

作者简介

张辰(1994—), 男, 博士生, 从事智能交通建模、控制与优化等研究, E-mail: yoyoke@sjtu.edu.cn;

徐云雯(1990—), 女, 副研究员, 博士, 从事预测控制、智能交通系等研究, E-mail: willing419@sjtu.edu.cn;

李德伟(1971—), 男, 教授, 博士生导师, 从事预测控制、鲁棒控制、滚动优化及其应用等研究, E-mail: dwli@sjtu.edu.cn.