

防治公交串车的近似动态规划法^{*}

何胜学

(上海理工大学 管理学院, 上海 200093)

摘要: 公交串车是公交系统运行失效的集中体现。为防治公交串车,建立了公交线路运行的近似动态规划模型,从而实现公交线路的动态自适应式控制。利用Q学习算法和基于人工神经网络近似的价值函数求解上述模型可以得到公交线路控制的最优策略,即基于系统状态确定站点滞留时间的状态价值函数。新的防治方法不仅可以利用仿真模型细致刻画公交线路的实际运行过程,而且可以动态整合在线和离线数据实现控制策略的动态优化调整。数值分析验证了新方法的有效性,与无控制情景相比,新方法不仅能防止公交串车发生,而且可有效降低公交车头时距的波动。

关键词: 智能交通; 公交串车; 驻车策略; 近似动态规划; Q学习

中图分类号: TP301 **文献标志码:** A **文章编号:** 1001-3695(2018)09-2624-04

doi: 10.3969/j.issn.1001-3695.2018.09.015

Approximate dynamic programming method of resisting bus bunching

He Shengxue

(School of Business, University of Shanghai for Science & Technology, Shanghai 200093, China)

Abstract: The operational failure of transit system reflects in the vicious bus bunching phenomena. To mitigate these negative influences, this paper formulated an approximate dynamic programming model to realize the dynamic adaptive control of a bus line. Using Q-learning approach and approximating the cost-to-go by artificial neural network representation, the optimal policy that was the optimal holding strategy for determining the proper slack times for a bus line could be found. New method can not only use simulation model to represent the reality in detail, but also has the feature of adaptive modification of holding policy by combining off-line learning and on-line application. Numerical experiment shows the efficiency of this new method. Compared with the no-control scenario, this new method can operate the bus system with no bunching and less variance of dynamic headway.

Key words: intelligent transportation; bus bunching; holding strategy; approximate dynamic programming; Q-learning

0 引言

公交串车现象是公交系统运行计划失效的集中体现^[1]。当发生串车时,乘客常需要候车较长时间。而经过长时间等待,往往会有两辆及以上公交车在短时间内相继到达。由于乘客一般会选择乘坐最先抵达的公交车,而该车往往已载有很多乘客,因此将加剧车内拥挤。而紧随其后的公交车由于上车乘客少,其大部分运能将被浪费。串车的直接危害包括乘客候车时间增加、线路运行计划失效和系统调度失灵等;其间接危害包括公交出行比重降低、交通拥堵加剧、能耗增加和环境污染恶化等。公交串车长期困扰公交运营企业,而如何有效防治公交串车则是公交运营管理研究领域长期关注的难点问题。现有的防治方法可大致分为起点发车控制、数值优化和自适应式控制三类。Osuna等人^[2]认为可以将环线公交的起点发车时间决策看做无限时域的Markov决策过程,通过控制发车间隔预防串车。Berrebi^[3]认为单点发车控制一般需要利用历史和预测数据,因此不适用经典Markov决策过程理论。安实等人^[4]认为滞站策略应被视为多阶段的随机决策过程,因此可以利用动态规划理论,以最小化车头时距波动为目标实现公交环线的单点控制。受限于单点控制,该类方法应用于较长线路时效果并不理想。

数值优化类方法常以最小化乘客出行时间为目标,通过求

解相应的数值优化模型来预防串车。由于串车的复杂随机特性,该方法一般被用于分析单控制点和仅有少量公交车的简单系统^[5,6]。Eberlein等人^[7]认为可以通过在滚动时域上优化乘客等待时间来实现串车防治。Delgado等人^[8]认为有效的串车防治有必要在滚动时域框架下整合站点滞留和登车限制两种策略。由于数值解析模型常简化或忽略系统的随机特征,所以该类方法的实用效果并不理想。

自适应式控制方法通过建立车辆运行状态与控制变量的直接关联关系实现串车防治。研究者分别从前向车头时距、后向车头时距和双向车头时距角度分析如何通过自适应式调节车速或设定驻站时间来保持车头时距的稳定性^[1,9-12]。但由于对影响站点停靠时间的随机因素考虑不足,上述方法很难在大扰动下预防串车。最近,一些研究者尝试利用计算智能技术,如机器学习算法、增强学习算法和模糊逻辑控制来实现串车防治^[13-15]。其中Luis等人^[13]考虑了站点滞留或跳站行车策略;陈春晓等人^[14,15]考虑了车速调节和驻站策略。考虑到这些新方法的各种简化假设和相关算法本身的效率,其实用效果有待进一步考证。

近似动态规划(approximate dynamic programming, ADP)是一种随机动态系统的有效控制方法。该方法具有广泛的适应性和较强的鲁棒性。本文将通过引入近似动态规划理论实现对公交串车现象的防治。在公交线路系统的近似动态规划模型建立之后,公交系统的控制就变为对驻站滞留控制策略不断

收稿日期: 2017-05-06; 修回日期: 2017-06-28 基金项目: 上海市自然科学基金资助项目(18ZR1426200/15ZR1429200); 上海市一流学科建设项目(S1201YLXK); 上海理工大学人文社科攀登重点项目(SK17PA02)

作者简介: 何胜学(1976-), 男, 陕西三原人, 副教授, 博士, 主要研究方向为交通控制、交通网络建模(lovellhe@126.com)。

的自适应式调整。现有实地数据可以提供系统的离线学习基础,而实时数据可作为系统的输入实现系统在线学习。将系统仿真嵌入本研究的理论模型,可以实现对公交系统更加细致的刻画。本研究的主要特色包括:a)首次构建了公交串车防治的近似动态规划模型;b)设计了可以有效求解新模型的基于人工神经网络(artificial neural network,ANN)近似价值函数的Q学习算法。

1 公交线路的近似动态规划模型

考虑一条具有 n_s 个站点的环形公交线路。线路上运行的公交车数目为 n_b 。本研究仅考虑最常见的公交车驻站滞留策略,即通过将载客后的车辆在站点滞留一定的时间来防治公交串车。研究假设系统的运行没有必须遵循的事先给定的车辆运行时刻计划或给定的车头时距。下面依次给出近似动态规划模型的六个主要组成部分。

1.1 控制行为集合

公交车在站点的滞留停车时间就是ADP理论中所说的控制行为。一般而言,控制行为集合应包括所有的可行滞留时间。如果以实数区间作为控制行为集,则势必引发维数灾难。因此只选择非负整数作为实际的控制行为。控制行为集合定义为 $A = \{0, 1, 2, \dots, a_u\}$, 其中 a_u 是最大的车辆可行滞留时间。公交线路系统演化过程中的不同阶段可能对应不同的控制行为集合,因此这里定义阶段 n 对应的控制行为集合为 A_n 。显然有 $A_n \subseteq A$ 成立。

1.2 系统状态集合

系统的状态对应于系统需要采取控制行为的时刻,需要全面涵盖系统运行的重要信息。利用系统状态量和系统实际采取的控制行为可以重现系统演化的全过程。为此设计了如下包含三个组成部分的系统状态量。对于任意一辆公交车 i , 需要考虑对应的两个特征量。第一个特征量 j_i 表示公交车 i 目前停靠或即将到达的公交站点。第二特征量 r_i 表示公交车 i 的等待激活时间,即从当前时刻到下一次对该公交车采取驻站滞留控制的时间长度。对任意公交站点 j , 需记录最近到达公交车的到达时间 p_j 。系统的状态量定义为 $s = (j_1, r_1, j_2, r_2, \dots, j_{n_b}, r_{n_b}, p_1, p_2, \dots, p_{n_s})^T$ 。令 $s_n = (s_{n,1}, s_{n,2}, \dots, s_{n,|s|})^T$ 为对应系统运行阶段 n 的系统状态量,其中 $s_{n,i}$ 表示其第 i 个元素。这里 $|s_n|$ 表示 s_n 包含的元素数目,其值等于 $2n_b + n_s$ 。

1.3 影响决策的随机信息

公交站点间的行车时间不确定性和乘客在站点上下车时间的不确定性是影响公交系统运行稳定性的主要随机因素。从公交站 i 到紧邻下游站点 j 的行车时间 $\hat{t}_{i,j}$ 可表示为

$$\hat{t}_{i,j} = t_{i,j} + \delta_{i,j} \quad (1)$$

其中: $t_{i,j}$ 为确定的行车时间均值; $\delta_{i,j}$ 为具有 0 均值和标准方差 $\sigma_{i,j}^2$ 的随机变量。假设 $\sigma_{i,j}^2$ 可以从实地调查数据推出。乘客在站点 i 的登车时间 \hat{t}_i 可表示为

$$\hat{t}_i = t_i + \delta_i \quad (2)$$

其中: t_i 为确定性时间部分,由到达乘客数目或与前车的车头时距决定; δ_i 是具有 0 均值和方差 σ_i^2 的随机变量。假设 σ_i^2 可由实地数据推出。为简化后续分析,假设在任意给定站点的乘客下车时间总小于乘客登车时间。

基于上述时间不确定性定义,本文定义影响决策的随机信息 $W_{n+1} = \{\delta_{i,j}^{n+1}, \delta_i^{n+1}\}$ 。这里上下标 $n+1$ 表示对应信息只有在系统的 $n+1$ 阶段开始后才有,且对应信息的实现时间为从系统阶段 n 到 $n+1$ 。

1.4 状态转移函数

公交线路系统运行的状态转移函数定义为

$$s_{n+1} = S^M(s_n, a_n, W_{n+1}) \quad (3)$$

上述抽象函数的具体逻辑执行过程如下:

a) 确定待激活车辆。待激活公交车 i 的等待激活时间满足 $r_i = \min\{r_1, r_2, r_3, \dots, r_{n_b}\}$ 。如果有两辆以上公交车满足最小等待激活时间条件,任取一辆即可。

b) 利用 W_{n+1} 的一个随机样本 $\omega_{n+1} = \{\delta_{j_i, j_i+1}^{n+1}, \delta_{j_i+1}^{n+1}\}$ 激活公交车 i 。更新待激活车辆的特征量 $r_i = r_i + a_n + \hat{t}_{j_i, j_i+1} + \hat{t}_{j_i+1}$ 和 $j_i = j_i + 1$ 。

c) 更新系统状态量。(a) 确定当前最小等待激活时间 $r_{\min} = \min_i\{r_i\}$; (b) 更新部分状态量 $r_i = r_i - r_{\min}$, $\forall i$ 和 $p_j = p_j - r_{\min}$, $\forall j$; (c) 利用更新后的 r_i 、 j_i 和 p_j 得到 s_{n+1} 。

1.5 系统绩效指标和控制行为代价函数

乘客等车时间和车辆运行速度是常用的两种公交系统绩效指标。考虑到现实中公交车速主要由路网交通特征和相应的基础设施决定,因此仅考虑乘客等车时间作为系统的绩效指标。等车时间可分为两种: c^1 定义为公交车到站前乘客的等车时间,而 c^2 定义为登车过程中乘客的车外等待上车时间。由于在本研究中系统运行基于仿真系统,上述两种等车时间可以在系统演化过程中直接取得。

等车时间也可以利用相关变量的随机分布特征计算得到。假设当前时间考虑在站点 j 激活公交车 i , 则公交车 i 与其紧邻的前车对于下游紧邻站点 j' 的车头时距 $H_{j,j'}^i$ 为

$$H_{j,j'}^i = a_n + \hat{t}_{j,j'} - p_{j'} \quad (4)$$

基于上述车头时距,等车时间 c^1 可依下式计算得到

$$c_j^1 = \frac{1}{2} \left(H_{j,j'}^i + \frac{\sigma_{j,j'}^2}{H_{j,j'}^i} \right) \vartheta_j H_{j,j'}^i \quad (5)$$

其中: ϑ_j 是公交站 j 的乘客到达率。同理, c^2 也可依据式(6)和(7)得出

$$\hat{t}_{j'}(\vartheta_j H_{j,j'}^i) = \psi \vartheta_j H_{j,j'}^i \quad (6)$$

$$c_j^2 = \frac{1}{2} \left(\hat{t}_{j'}(\vartheta_j H_{j,j'}^i) + \frac{\hat{\sigma}_{j'}^2}{\hat{t}_{j'}(\vartheta_j H_{j,j'}^i)} \right) \vartheta_j H_{j,j'}^i \quad (7)$$

其中: ψ 表示一个乘客的平均登车时间。式(4)~(7)的物理含义可参阅文献[16]。

本研究考虑两种系统状态稳定性的度量指标。受已有研究[1]启发,可定义一个期望系统车头时距。假设公交线路长度为 l , 公交车平均行驶速度为 \bar{v} , 则期望系统车头时距可由下式计算得到

$$\bar{H} = (l + \sum_i w_i \bar{v} + \sum_j t_j \bar{v}) / (\bar{v} n_b) \quad (8)$$

其中: w_i 是车辆在交叉口 i 的期望等待时间; t_j 是公交车在站点 j 的期望停留时间。

另一个系统车头时距为动态环线车头时距。定义如下:

$$\bar{H} = \sum_i h_i / n_b \quad (9)$$

其中: h_i 表示公交车 i 与其紧邻前车的车头时距。

车头时距波动是一种有效的系统动态不稳定性度量指标。利用式(8)和(9),可定义针对不同系统阶段的两个系统控制行为代价函数:

$$c_a = \sum_i (h_i - \bar{H})^2 \quad (10)$$

或者

$$c_a = \sum_i (h_i - \hat{H})^2 \quad (11)$$

在随后的数值分析中,本文将比较两种控制行为代价函数的应用效果。

1.6 控制的最优策略

近似动态规划中的策略指的是根据系统状态选择相应控

制行为的规则,也可以将其视为一个函数。而确定最优策略就是在所有可能的函数中选择使得目标最优的一个,因此近似动态规划的本质可以视为泛函优化。设定 $\pi \in \Pi$ 是一个一般性策略, Π 为可行策略集合。系统状态与控制行为的函数关系可表示为

$$a = A^\pi(s) \quad (12)$$

如果将系统的绩效指标直接作为优化目标,则控制行为代价函数也可表示为

$$C(s_n, A^\pi(s_n), W_{n+1}) = (c^1(s_n, A^\pi(s_n), W_{n+1}) + c^2(s_n, A^\pi(s_n), W_{n+1})) \quad (13)$$

其中: $c^1(s_n, A^\pi(s_n), W_{n+1})$ 和 $c^2(s_n, A^\pi(s_n), W_{n+1})$ 是对应系统从状态 s_n 演化到状态 s_{n+1} 的乘客等待时间。当然也可以利用式(10)和(11)给出的控制行为代价函数建立目标函数。大量的数值分析得出利用后者建立的行为代价函数更为有效,可以大大提高算法的收敛速度。后面将统一用 $C(s_n, A^\pi(s_n), W_{n+1})$ 表示代价函数,不再细分上述差别。

如果系统的总阶段数有限,则可用 $\min_{\pi \in \Pi} \mathbb{E}[\sum_{n=0}^{N-1} C(s_n, A^\pi(s_n), W_{n+1})]$ 作为系统优化目标。这里 $\mathbb{E}(\cdot)$ 表示求随机变量的期望值。有无限个阶段的系统则可考虑滚动时域的第 n 阶段优化目标 $\min_{\pi \in \Pi} \mathbb{E}[\sum_{n=0}^{\infty} C(s_n, A^\pi(s_n), W_{n+1})]$ 。但是由于具有无限个阶段的系统优化目标值一般很大或趋于无穷,所以可以考虑随着时间推移,控制行为的代价影响减弱。对于任意给定的阶段 l 可用下面带折扣因子 κ 的加和式作为优化目标:

$$\min_{\pi \in \Pi} \mathbb{E}[\sum_{n=l}^{\infty} \kappa^{n-l} C(s_n, A^\pi(s_n), W_{n+1})] \quad (14)$$

直接求解以式(14)为目标函数,以其他系统运行条件为约束的优化问题非常困难。因此常用的方法是根据 Bellman 方程,迭代求解一个均衡条件下的价值函数。设公交系统处于状态 s 的价值函数为 $V(s)$ 。依据 Bellman 方程,在无折扣条件下, $V(s)$ 需满足

$$V^*(s_n) = \mathbb{E}[-C(s_n, A^*(s_n), W_{n+1}) + V^*(s_{n+1} | s_n, A^*(s_n), W_{n+1})] \quad (15)$$

当考虑带有折扣因子的优化目标式(14)时, $V(s)$ 应满足

$$V^*(s_n) = \mathbb{E}[-C(s_n, A^*(s_n), W_{n+1}) + \kappa V^*(s_{n+1} | s_n, A^*(s_n), W_{n+1})] \quad (16)$$

实际应用中,折扣因子 κ 常被收敛序列 α_n (满足 $0 \leq \alpha_n < 1$, $\sum_n \alpha_n < +\infty$, $\forall n$) 代替。依据经典近似动态规划理论,上述收敛序列 α_n 可以确保优化目标的收敛性,从而保证后续算法的收敛性。

2 基于 ANN 近似价值函数的 Q 学习算法

考虑到神经网络能实现对各种函数关系的良好拟合,本文将利用 ANN 对算法中的价值函数加以近似表示。在 Q 学习算法中价值函数即为 Q 函数。状态价值函数 $V(s)$ 与 Q 函数 $Q(s, a)$ 具有如下基本关系:

$$V(s) = \max_a Q(s, a) \quad (17)$$

在下面给出的算法中, n 是当前阶段的序列号, m 是对系统进行仿真模拟的序号。 N 和 M 分别是总的阶段数目和模拟仿真总次数。 θ 是待定的人工神经网络连接权重。针对公交车串车进行防治的 Q 学习算法的求解具体步骤如下:

a) 初始化。

(a) 对所有的系统状态 s_n , 给出价值函数 $\bar{Q}_n^0(s_n, a_n)$ 的近似值以及控制行为 $a_n \in A_n, n = \{0, 1, \dots, N-1\}$ 。

(b) 令 $m = 1$ 。

(c) 初始化 s_0^1 。

b) 选择一个随机信息的样本路径 ω^m 。

c) 将下面操作步骤依 $n = 0, 1, \dots, N-1$ 加以循环迭代。

(a) 利用 ε 贪婪规则确定控制行为。以概率 ε , 从控制行为集 A_n 中随机选择行为 a^m 。而以概率 $1 - \varepsilon$, 利用公式 $a_n^m \in \operatorname{argmax}_{a_n \in A_n} \bar{Q}_n^{m-1}(s_n^m, a_n, \theta)$ 选择行为 a^m 。

(b) 对信息 $W_{n+1}^m = W_{n+1}(\omega^m)$ 进行取样, 并计算下一状态 $s_{n+1}^m = S^M(s_n^m, a_n^m, W_{n+1}^m)$ 。

(c) 计算 $\hat{q}_n^m = -C(s_n^m, a_n^m) + \gamma \max_{a_{n+1} \in A_{n+1}} \bar{Q}_{n+1}^{m-1}(s_{n+1}^m, a_{n+1}, \theta)$ 。

(d) 更新 \bar{Q}_n^m 和 θ 。

首先, 依据下式对 θ 进行更新:

$$\theta := \theta + \lambda \nabla \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta) (\hat{q}_n^m - \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta))$$

也可按如下方式对 θ 加以更新:

$$\theta := \theta + \lambda \nabla \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta) (-C(s_n^m, a_n^m) + \gamma \bar{Q}_{n+1}^{m-1}(s_{n+1}^m, a_{n+1}, \theta) - \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta))$$

在上述操作中, 控制行为 a_{n+1} 可由 ε 贪婪规则确定。

其次, 基于更新后的参数 θ , 对 Q 因子进行如下更新:

$$\bar{Q}_n^m(s_n^m, a_n^m) = (1 - \alpha_{m-1}) \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta) + \alpha_{m-1} \hat{q}_n^m$$

d) 令 $m := m + 1$ 。如果 $m \leq M$, 转步骤 b)。

e) 给出最终的 Q 因子 $(\bar{Q}_n^M)_{n=1}^{N-1}$ 和 θ 。

在一定意义上, 上述算法的目的就是最小化 θ 的函数 $f(\theta) = (1/2) (\hat{q}_n^m - \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta))^2$ 。函数 $f(\theta)$ 的负梯度方向为 $\nabla \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta) (\hat{q}_n^m - \bar{Q}_n^{m-1}(s_n^m, a_n^m, \theta))$ 。这里 θ 作为自变量, \hat{q}_n^m 为给定值。按照经典非线性规划理论, 在 θ 的可行域内, 如果当前的 θ 非局部或全局最小值, 且步长足够小, 目标函数 $f(\theta)$ 的值将沿负梯度方向下降。按照上述分析, 算法中的参数 λ 将相当于搜索步长。因此可以通过试错法确定 λ 的合理取值。与经典非线性规划理论主要针对确定性系统不同, Q 学习算法针对的是随机系统的实时自适应控制。上述 Q 学习算法可以利用随机信息的样本路径 ω^m 将已有信息和实时信息结合, 并利用贪婪规则实现对已有优化结果的利用和对未知 Q 函数空间的探索, 从而最终达到维持系统运行长期稳定的目标。与经典非线性规划算法一般会逐渐收敛于一个局部最优解不同, Q 学习算法的优化结果一般具有全局最优特征。

3 数值算例

考虑一条具有 18 个站点和 5 辆公交车的环形公交线路。假设车辆的平均巡游速度为 30 km/h, 仿真的系统时长为 2 h。表 1 给出了公交车的初始信息, 其中缩写“OWTA”代表初始等待激活时间(original waiting time to be activated)。车辆载客能力为车辆任意时刻可载运的最大乘客数。表 2 给出了站点的基本信息。表 2 中乘客到达率的单位为人/min, 而站间距指两个相邻站点间线路长度, 在表中用前面的一个站点作为标示。

表 1 公交车初始信息

初始信息	车辆编号				
	1	2	3	4	5
载客能力	100	80	80	100	80
初始站点	1	4	8	11	15
OWTA/s	20	0	40	30	50

表 3 给出了公交线分段所包含路段的长度和交叉口交通控制信息。这里的公交线分段指的是连接两个相邻公交站点的公交线路部分, 而这里的路段则是由公交站点和交叉口分割得到公交线路部分。算例中假设交叉口的信号控制均为定时控制。四个控制参数中, 第一个是红灯时间(s); 第二个是绿灯时间(s); 第三个参数表示初始相位已经持续的时长(s); 第四个参数指示具体相位, 1 表示红灯相位, 2 表示绿灯相位。这

里忽略了黄灯时间。未出现在表3中的公交线分段仅包含一条路段,相关信息如表2所示。

表2 乘客到达率和站间距离

参数	站点序号																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
乘客到达率	5	4	5	1	1	1	5	1	2	4	5	2	1	2	5	1	1	2
站间距离/m	600	500	600	600	600	500	600	800	600	750	600	700	600	500	600	500	500	600

表3 路段长度和信号交叉口控制参数

参数	公交线分段序号														
	1		4		8		10		12		13		16		
包含的路段长度/m	200	400	250	350	300	500	300	350	300	400	300	300	200	200	100
交叉口控制参数	40,50,20,2		40,30,20,2		40,35,20,2		30,45,20,2		30,40,20,2		40,30,20,2		40,55,20,2		30,35,20,2

用于近似价值函数的人工神经网络为包含两个隐含层的多层感知机模型。输入层包含 $29(2 \times n_h + n_s + 1)$ 个节点,输出层包含1个节点,而两个隐含层均包含5个节点。参数 β 等于0.5的正切Sigmoid函数作为神经网络的激活函数。神经网络连线的初始权重从区间 $(-2,2)$ 中随机选取。

两个连续控制行为 a_k 和 a_{k+1} 的时间差设为2s。算法总的仿真次数为300次。考虑11个差异化的控制行为。 a_0 表示无控制,即不对车辆进行站点滞留。令 $\varepsilon=0.8$,折扣因子 $\gamma=0.3$,初始步长 $\lambda=0.01$ 。考虑两类乘客,第一类的平均登车时间为4s,平均下车时间为2s;第二类乘客的平均登车时间为1s,下车时间为0.5s。第一类乘客与第二乘客的数量比为1:10。

图1给出了一个典型的无控制情景下的公交车运行时空轨迹。从图1中可以看到串车现象的发生。第一次串车出现在仿真时间3214s的站点8,串车现象持续的时间很久。大量的仿真实验证明在2h的观测期内该线路的串车现象几乎不可避免。

图2给出了一个典型的利用期望系统车头时距进行优化的车辆运行时空轨迹仿真结果。系统的期望系统车头时距大约为360s。大量实验证明利用期望系统车头时距的方法可有效防止公交串车的发生,但是控制系统采用的滞留车辆时间一般会取模型允许的最大值,因此实际应用时乘客的直接感受将较差。

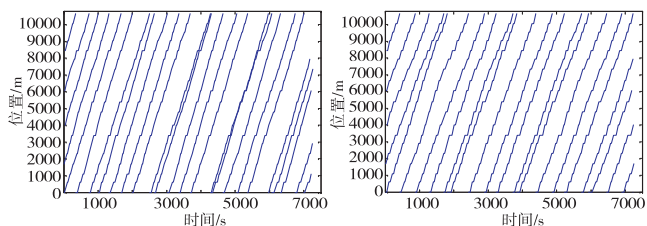


图1 无控制情景下公交车运行轨迹

图2 利用期望系统车头时距的控制情景下公交车运行轨迹

图3给出了利用动态环线车头时距的一个典型车辆运行轨迹仿真结果。可以看出该方法可以使车辆的运行轨迹在时空图中的分布更加均匀。图4给出了对应图3的系统控制滞留时间序列。可以看出该方法所采用的滞留时间也是比较均匀地分布在可行区间 $[0,20]$ 上。对应的系统总等车时间为344056s。对一次控制而言,平均的等车时间为9773s,其标准方差为15455s。如果假设公交车的载客能力为100人,且车辆总是处于半载状态,则每一秒的滞留时间将被50个乘客体验。这种情况下,总的滞留时间只占总等车时间的约5%。

图5给出了对应图3的乘客登车时间序列。累积的总登车时间为6242s。对于一次控制行为而言,平均的登车时间为17.7s,其标准方差为12.2s。从图5可以看出,乘客的站点到达是完全随机的。

通过对过去经验的学习总结,ANN近似可得到对系统未

来运行指标的预测值,即ANN近似价值函数值。而实际仿真所得价值函数值是利用仿真系统在给定随机信息样本条件下生成的系统运行指标值。图6给出了ANN近似和实际仿真所得价值函数值的差值占仿真所得价值函数值的比值。图6数据展现了利用ANN来近似价值函数的效率。该比值的变化范围较小,且随着系统运行时间增加,上述比值的变化范围逐渐缩小。这说明随着输入数据的不断增加,利用ANN近似实际价值函数的效果越来越好。但考虑到系统本身的随机动态变化特征,上述比值将呈现一个持续波动的趋势,而不是收敛于0。可以预期当系统随机因素出现小概率事件时,比值的波动幅度会有所增加。

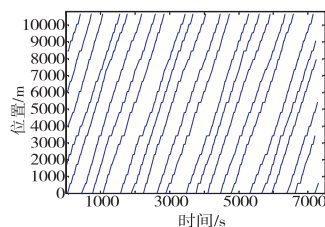


图3 利用动态环线车头时距的控制情景下公交车运行轨迹

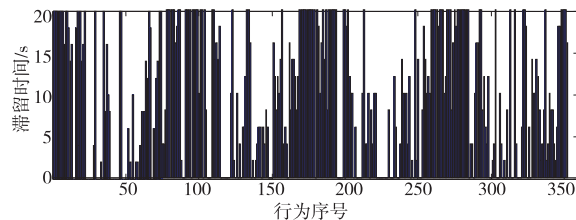


图4 最后一轮仿真所得的最优滞留时间序列

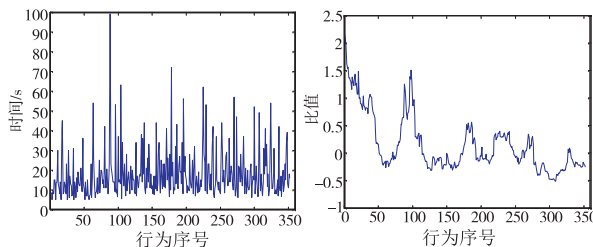


图5 登车时间序列

图6 ANN近似和系统仿真所得价值函数的差与仿真所得价值函数值的比值

图7给出了ANN一个指定连接的权重的更新变化过程。可以看出该连接权重的变化处于区间 $[84700,84900]$ 。与连接权重最终选取的具体值84800相比,该变化范围还是很小的。这种连接权重的波动体现了Q学习算法不断更新价值函数ANN近似的努力。权重波动幅度的变化趋势与图6中的比值变化基本一致。以行为序号150为分界线,行为序号小于150的比值和连接权重波动均较大,而序号大于150的比值和连接权重波动较小。图6和7的数据变化趋势说明Q学习算法具有较高的可靠性和较强的鲁棒性。

(下转第2658页)

进行了可行性和有效性验证。

参考文献:

- [1] 相亮亮. 无人机在军事中的应用与发展[J]. 科技展望, 2016, 26(14): 292.
- [2] 龙威林. 无人机的发展与应用[J]. 产业与科技论坛, 2014, 13(8): 68-69.
- [3] 刘一鸣. 当无人机驶入“禁飞区”[J]. 中国企业家, 2015(7): 116-117.
- [4] Zhang Jun. The research of UVA technology application development and regulation[J]. *Modern Science & Technology of Telecommunications*, 2016, 46(3): 69-71, 78.
- [5] Ayhan S, Samet H. Aircraft trajectory prediction made easy with predictive analytics[C]//Proc of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2016: 21-30.
- [6] Schuster W. Trajectory prediction for future air traffic management-complex manoeuvres and taxiing[J]. *Aeronautical Journal*, 2015, 119(1212): 121-143.
- [7] 郝杰, 任静. 高超声速飞行器交互式多模型跟踪算法仿真[J]. 计算机技术与应用, 2015, 25(2): 204-206.
- [8] 翟岱亮, 雷虎民, 李炯, 等. 基于自适应 IMM 的高超声速飞行器轨迹预测[J]. 航空学报, 2016, 37(11): 3466-3475.
- [9] 徐庆飞, 张新, 李卫民. 二维空间中目标轨迹预测算法研究与分析[J]. 航空电子技术, 2012, 43(1): 10-14.
- [10] Zernetsch S, Kohnen S, Goldhammer M, et al. Trajectory prediction of cyclists using a physical model and an artificial neural network[C]//Proc of IEEE Intelligent Vehicles Symposium. Piscataway, NJ: IEEE Press, 2016: 833-838.
- [11] 徐肖豪, 杨国庆, 刘建国. 空中飞行轨迹预测算法的比较研究[J]. 中国民航大学学报, 2001, 19(6): 1-6.
- [12] 王涛波, 黄宝军. 基于改进卡尔曼滤波的四维飞行轨迹预测模型[J]. 计算机应用, 2014, 34(6): 1812-1815.

- [13] Qi Le, Zheng Zhongyi. Trajectory prediction of vessels based on data mining and machine learning[J]. *Journal of Digital Information Management*, 2016, 14(1): 33-40.
- [14] Feng Zhenni, Zhu Yanmin. A survey on trajectory data mining: techniques and applications[J]. *IEEE Access*, 2016, 4: 2056-2067.
- [15] Xie Yu, Liu Luhua, Liu Jun, et al. Rapid generation of entry trajectories with waypoint and no-fly zone constraints[J]. *Acta Astronautica*, 2012, 77(8): 167-181.
- [16] Liang Zixuan, Liu Siyuan, Li Qingdong, et al. Lateral entry guidance with no-fly zone constraint[J]. *Aerospace Science & Technology*, 2016, 60(1): 39-47.
- [17] Temizer S, Kochenderfer M, Kaelbling L, et al. Collision avoidance for unmanned aircraft using Markov decision processes[C]//Proc of AIAA Guidance, Navigation, and Control Conference. Piscataway, NJ: IEEE Press, 2010: 1-22.
- [18] Kim K Y, Park J W, Tahk M J. UAV collision avoidance using probabilistic method in 3-D[C]//Proc of International Conference on Control, Automation and Systems. Piscataway, NJ: IEEE Press, 2007: 826-829.
- [19] Chen Mingjing, Fang Yuanmin, Chen Jie. Fitting of circular curve based on least square method and iterative method[J]. *Science of Surveying and Mapping*, 2016, 41(1): 194-197, 202.
- [20] 陈伟杰. 最小二乘法原理及其在实验曲线拟合中的应用分析[J]. 辽宁科技学院学报, 2014, 16(4): 33-34.
- [21] 陈岚峰, 杨静瑜, 崔崧, 等. 基于 MATLAB 的最小二乘基线拟合仿真研究[J]. 沈阳师范大学学报: 自然科学版, 2014, 32(1): 75-79.
- [22] 顾天奇, 张雷, 冀世军, 等. 封闭离散点的曲线拟合方法[J]. 吉林大学学报: 工学版, 2015, 45(2): 437-441.
- [23] 李蓓蓓. 多次自适应最小二乘曲线拟合方法及其应用[D]. 荆州: 长江大学, 2014.
- [24] 陈明, 张丰, 杜震洪, 等. 环状分布平面点集的凸包快速生成算法[J]. 上海交通大学学报, 2014, 48(5): 658-662.
- [25] 刘斌, 王涛. 一种高效的平面点集凸包递归算法[J]. 自动化学报, 2012, 38(8): 1375-1379.

(上接第 2627 页)

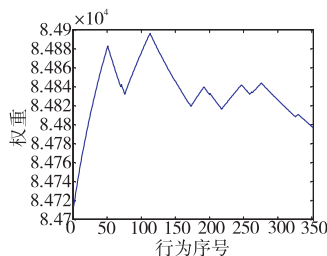


图7 ANN中指定连接的权重变化趋势

4 结束语

针对如何防治公交串车现象, 基于近似动态规划理论建立了公交线路系统的动态自适应式控制系统。该系统可以动态地决策是否对公交车进行站点滞留控制, 并确定具体的滞留时间长短, 从而达到预防串车发生、增强系统稳定性和提升乘客体验的效果。设计的基于 ANN 近似价值函数的 Q 学习算法具有良好的收敛性, 可以有效地利用在线和离线数据。通过定义两种不同的系统车头时距, 实现对系统整体稳定性的有效度量。研究表明动态环线车头时距比期望系统车头时距的应用效果更佳。本研究提出的方法能用于现实公交线路的动态协调控制, 提高线路服务水平。后续研究可以进一步考虑多线路相互影响下的串车防治, 以及地面其他交通方式对公交车行驶的影响。

参考文献:

- [1] He Shengxue. An anti-bunching strategy to improve bus schedule and headway reliability by making use of the available accurate information[J]. *Computers & Industrial Engineering*, 2015, 85(1): 17-32.
- [2] Osuna E E, Newell G F. Control strategies for an idealized bus system[J]. *Transportation Science*, 1972, 6(1): 52-71.
- [3] Berrebi S J. A real-time bus dispatching policy to minimize headway

- variance[D]. Atlanta: Georgia Institute of Technology, 2014.
- [4] 安实, 张昕明, 王健. 基于随机决策的公交车辆滞站策略[J]. 科学技术与工程, 2016, 16(28): 115-119, 125.
- [5] Hickman M D. An analytic stochastic model for the transit vehicle holding problem[J]. *Transportation Science*, 2001, 35(3): 215-237.
- [6] Zhao Jiamin, Dessouky M, Bukkapatnam S. Optimal slack time for schedule-based transit operations[J]. *Transportation Science*, 2006, 40(4): 529-539.
- [7] Eberlein X J, Wilson N H M, Bernstein D. The holding problem with real-time information available[J]. *Transportation Science*, 2001, 35(1): 1-18.
- [8] Delgado F, Muñoz J C, Giesen R. How much can holding and/or limiting boarding improve transit performance? [J]. *Transportation Research, Part B*, 2012, 46(9): 1202-1217.
- [9] Daganzo C F. A headway-based approach to eliminate bus bunching: systematic analysis and comparisons[J]. *Transportation Research, Part B*, 2009, 43(10): 913-921.
- [10] Xuan Yiguang, Argote J, Daganzo C F. Dynamic bus holding strategies for schedule reliability: optimal linear control and performance analysis[J]. *Transportation Research, Part B*, 2011, 45(10): 1831-1845.
- [11] Bartholdi III J J, Eisenstein D D. A self-coordinating bus route to resist bus bunching[J]. *Transportation Research, Part B*, 2012, 46(4): 481-491.
- [12] Daganzo C F, Pilachowski J. Reducing bunching with bus-to-bus cooperation[J]. *Transportation Research, Part B*, 2011, 45(1): 267-277.
- [13] Luis M M, Oded C, Joao G, et al. An online learning approach to eliminate bus bunching in real-time[J]. *Applied Soft Computing*, 2016, 47(7): 460-482.
- [14] 陈春晓, 陈治亚, 陈维亚. 基于多智能体增强学习的公交驻站控制方法[J]. 计算机工程与应用, 2015, 51(17): 8-13, 27.
- [15] 陈春晓, 陈治亚, 陈维亚. 基于模糊逻辑的单线路公交实时控制方法[J]. 公路交通科技, 2016, 33(9): 141-147.
- [16] Daganzo C F. Passenger waiting time: advertised schedules[M]//Fundamentals of Transportation and Traffic Operations. New York: Elsevier, 1997: 291-292.