DOI:10.11918/202502033

基于 Actor-Critic 网络深度强化学习的定制公交线网优化

李 冰,周晓倩,轩 华

(郑州大学 管理学院,郑州 450001)

摘 要:为了研究一类考虑多车种可用的多对多定制公交线网优化问题,以乘客出行成本和公交运营成本最小为目标,考虑乘降点服务、乘降点数量、线路总长度以及公交车队规模等限制,构建多对多定制公交线网优化双层混合整数规划模型。传统方法在求解大规模数学模型方面效率不高,提出一种基于 Actor-Critic 双重神经网络的深度强化学习方法。该方法将定制公交视为智能体,利用无监督聚类建立智能体学习环境,构造智能体状态-动作-奖励机制,设计 Actor-Critic 双重神经网络用于乘降点选择与评价,构建基于 Actor-Critic 双重网络的深度强化学习过程。选取郑州市高新区-郑东新区公交乘降点分布与居民出行数据搭建实验场景,对学习率、折扣因子、批尺寸和神经元数等超参数进行优化调试,利用双重网络深度强化学习过程求解模型,得到多对多定制公交线网方案,其乘降点分配、行驶距离、上座率与车型安排都较为合理。与基于 Actor 单重网络的深度强化学习进行结果对比,算法收敛速度提高了 28.57%,奖励值增加了 13.65%;进一步地,与启发式算法对比也验证了本文提出的方法可以快速高效完成模型求解,合理确定定制公交乘降点位置和具体线路走向。

关键词:Actor-Critic 网络;定制公交;乘降点;人工神经网络;深度强化学习

中图分类号:N945;F530 **文献标志码**:A **文章编号**:0367-6234(2026)07-0085-12

Optimization of customized bus route based on deep reinforcement learning with Actor-Critic neural network

Li Bing, Zhou Xiaoqian, Xuan Hua

(School of Management, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The optimization of customized bus route with multiple boarding sites and alighting sites considering heterogeneous fleet is studied. Aiming at minimizing passenger trip cost and bus operating cost, the problem is formulated as a mathematical programming model where some constraints indicating the boarding and alighting mode, site number, the length of bus route, the number of customized buses are considered. The traditional methods are not efficient in solving large-scale mathematical models. So a deep reinforcement learning procedure based on Actor-Critic dual neural network is proposed. Here the customized bus is regarded as agent. The learning environment of the agent is established by the unsupervised clustering approach. And then the state space-action space-reward function of the agent is presented and a Actor-Critic double neural network is given to select and evaluate boarding sites and alighting site. So a deep reinforcement learning algorithm based on Actor-Critic dual neural network is obtained. The coordinate data of boarding-alighting sites and resident trip is collected together to set up testing experiment. These hyper-parameters, which is composed of learning rate, discount factor, batch size and number of neurons, are debugged by New Actor neural network and Old Actor neural network. The customized bus route scheme based on multiple boarding sites and alighting sites can be obtained by the proposed deep reinforcement learning algorithm. The performance measures, which consist of the layout of boarding and alighting sites, customized bus route distance, bus occupancy rate and bus type allocation, verify the rationality of the customized bus route scheme. The deep reinforcement learning algorithm based on single Actor neural network is used as a comparison testing. The convergence speed of proposed algorithm is improved by 28.57% and the reward value is increased by 13.65%. Furthermore, compared with the heuristic algorithms, this paper also verifies that the proposed method can quickly and efficiently complete the model solution, and reasonably determine the location of the customized bus stops and the specific route direction.

Keywords: Actor-Critic neural network; customized bus; boarding and alighting site; artificial neural network; deep reinforcement learning

收稿日期:2025-02-20;**录用日期**:2025-04-02;**网络首发日期**:2025-08-08
网络首发地址:https://link.cnki.net/urlid/23.1235.T.20250808.1217.003
基金项目:国家社会科学基金(24BJY113)
作者简介:李 冰(1976—),男,教授,博士生导师
通信作者:周晓倩,xiaoqianzhou21@163.com

随着城市公共交通优先发展战略的实施,城市公交系统快速发展,公交服务水平显著提升,公交出行成为居民最主要的出行方式之一。然而,居民的多样化出行需求与常规公交单一服务模式之间的矛盾越来越突出,制约了城市公交系统进一步提升。在“互联网+交通”背景下,出现了以乘客出行需求为导向,提供个性化服务的定制公交,根据客流预约出行需求选择乘降点,将区域预约需求转化为线路需求,从而给出更为合理的公交线路,满足乘客出行需求。对于跨区域长距离城市通勤需求,定制公交服务具有良好的需求响应效果。但是定制公交的实际运营存在上座率不高、运行成本高、运行时间长等问题,使得中国城市定制公交系统发展缓慢。如何依据乘客出行需求,合理布局乘降点位置、有序安排公交线路走向亟待深入研究。

关于定制公交的研究起源于“预约+共享”这一概念,早期研究主要集中在理论探讨层面。Liu和Ceder[1]首次使用定制公交这一概念,根据开通定制公交的30个城市数据,总结了定制公交的背景、分类、规划运营等方面内容。Wang等[2]研究了定制公交系统在整个城市公共交通系统中的定位,即定制公交的主要服务群体、市场定位、系统建设,Wang等[3]研究了定制公交服务的空间依赖性问题。

引入量化技术是解决定制公交问题的有效方法。在面向多上车点单下车点的多对一定制公交线路优化问题研究方面,Ellegood等[4]研究一类校车接送的多对一定制公交路线,采用连续近似模型为多个学校运送学生路线提供有利条件;安久煜等[5]研究考虑回程空载的资源浪费的接驳性定制公交线网调度优化模型;王正武[6]提出接驳公交基于关键点的路径优化方法。在面向多上车点多下车点的多对多定制公交线路优化方面,Guo等[7]研究定制公交路线设计,构建混合整数规划模型并用遗传和分支定界算法求解对比。马昌喜等[8]以现有多个上下车站点,构建突发公共卫生事件下的应急定制公交线路优化方法。现有研究集中关注乘降点数量而未考虑基于乘客出行需求的乘降点位置规划,但定制公交线路受乘降点数量和位置等因素影响。

目前开展的定制公交线网优化研究主要从乘客便利度或公交公司运营效率角度展开。马昌喜等[8]以定制公交总运行时长最短为目标构建模型;王超等[9]考虑总运营里程最小建立模型;Li等[10]考虑运营车辆数量最少构建目标函数。只有少数学者研究了考虑乘客和企业的优化目标,Wang等[11]以乘车消耗和公交运营成本最小为目标,提出两阶段优化模型进行求解。陈汐等[12]以最小化乘客出行成本、车辆运营成本为优化目标;Huang等[13]在最大服务率和最小运营成本目标条件下研究定制公交线路优化问题。上述研究均未考虑多种车型,仅探讨了优化目标对定制公交线网优化的影响。也有学者研究了多车型的定制公交单目标优化问题,胡郁葱等[14]以总成本最小构建了考虑多起终点、多车型的定制公交线路规划模型,并利用遗传算法改进编码规则;何民等[15]采用不同车型来体现乘客舒适性,构建考虑响应式定制公交运营商成本最小和出行可靠性、舒适性要求的公交线路设计模型。随着人工智能的高速发展,利用人工智能算法来解决线网优化问题成为进一步的发展方向,Wang等[16]研究考虑跨区域的路径规划,提出基于Q-learning算法的定制公交路线优化方法。彭理群等[17]提出一种跨区域定制公交的搭乘方案,通过改进Q-learning算法对公交线路进行优化。现有研究利用启发式算法或强化学习求解优化模型,但是面对复杂的环境和高维的数据,启发式算法或强化学习算法求解效率不理想。表1根据乘降点数量、乘降点规划、优化目标、车型以及求解算法的不同,罗列了相关研究问题的结果。

由表1可知,已有研究工作在开展线网优化工作中,在公交车型多样化、乘降点规划调整等方面考虑不多,且在研究方法设计方面多采用商用求解器、启发式算法来完成求解任务。面对市场的竞争化机制和乘客实际使用需求,运营公司需要建立定制公交服务的快速需求响应机制[18],但考虑到定制公交线路规划是一个动态实时决策过程,会存在乘降点位置变化、乘客出行需求异动、公交车型变动等诸多可变因素干扰,为快速需求响应机制带来不小的挑战。显然,已有研究所采用的方法无法有效应对需求响应式定制公交的实时应答要求。为弥补上述研究的不足,本文将在以下方面进行拓展:一是研究考虑多车型定制公交车队;二是综合考虑乘客出行成本与公交运营成本;三是依据乘客出行需求完成乘降点实时规划;四是提出双重神经网络深度强化学习求解方法。

表 1　相关研究问题结果展示

Tab. 1　Results for the research problems considered

文献	乘降点数量		乘降点规划		优化目标		车型	量化技术
	多上车点单下车点	多上车点多下车点	基于历史经验	基于出行需求	单目标优化	多目标优化		
[6]	√		√		√		单	遗传算法
[7]		√	√			√	单	遗传算法
[8]		√	√		√		单	遗传算法
[9]		√		√	√		单	遗传算法
[10]		√	√		√		单	分支定界算法
[11]		√	√			√	单	分支定界算法
[12]		√	√			√	单	两阶段启发式
[13]		√	√		√		单	邻域搜索算法
[14]		√		√	√		多	遗传算法
[15]		√	√		√		多	遗传算法
[16]		√	√			√	单	强化学习
[17]		√	√		√		单	强化学习
本文		√		√		√	多	深度强化学习

1　问题建模

1.1　问题描述

多对多定制公交线网优化问题可描述为：运营公司根据乘客的出行需求划分出上车区域和下车区域，上车区域和下车区域之间存在多个乘客行程。每辆定制公交从上车区域某个上车点出发，并继续前往其他上车点搭载乘客，然后再将乘客一一送至下车区域的相应下车点，直至完成车上所有乘客出行需求。公交运营公司需要根据乘客出行需求，在上车区域和下车区域合理安排上下车站点位置，设计不同车型定制公交的运行线路，在满足公交载客量以及线路开行条件下，将乘客送达目的地，并最小化乘客出行成本和公交企业运营成本。

为简化模型，本文作出以下假设：首先，每辆车只能服务上车点的一个行程（一个上车点可能存在多个行程）；其次，定制公交车辆在上车区域接完所有乘客后，才能前往下车区域将乘客送到相应下车点；最后，上车区域内上车点乘客在只上不下，下车区域内下车点乘客只下不上。

1.2　符号说明

为构建模型，引入集合、参数与变量见表 2。

表 2　符号说明

Tab. 2　Notation and definitions

类别	符号	定义
集合	O	上车区域乘客上车点集合，记为 $O=\{o \mid o=1,\cdots,\tilde{o}\}$，$\tilde{o}$ 为乘客上车点总数
	D	下车区域乘客下车点集合，记为 $D=\{d \mid d=1,\cdots,\tilde{d}\}$，$\tilde{d}$ 为乘客下车点总数
	V	乘降点集合，记为 $V=O\cup D$，O 为乘客上车点集合，D 为乘客下车点集合
	ψ	乘客的 OD 行程集合，其中，单个 od 行程记为 (o,d)
	W	定制公交车辆集合，记为 $W=\{w \mid w=1,\cdots,\tilde{w}\}$，$\tilde{w}$ 为定制公交车辆总数
	M	定制公交车辆类型集合，记为 $M=\{m \mid m=1,\cdots,\tilde{m}\}$，$\tilde{m}$ 为定制公交车型总数
参数	n_{od}	单个 od 行程的乘客人数，$o\in O$，$d\in D$。单位：人
	k	定制公交可用车辆数。单位：辆
	p_{ij}	乘降点 i 到乘降点 j 的距离，$i,j\in V$。单位：km
	ξ_1^m	m 类型定制公交车辆的单次发车成本，$m\in M$。单位：元/次
	ξ_2^m	m 类型定制公交车辆的单位距离运行成本。单位：元/km
	c	单位乘客乘车时间成本，单位：元/（人·h）
	φ	单条定制公交线路的乘降点最大访问数量。单位：个
	μ	定制公交车辆开行最小上座率，取 75%
	f_m	m 型定制公交车辆的最大载客量，$m\in M$。单位：人
	e	定制公交车辆行驶速度。单位：km/h
	l	定制公交续航里程。单位：km

表 2(续)

类别	符号	定义
决策变量	x^{mw}	m 型公交车辆 w 投入运营时取 1,否则取 0。$w \in W, m \in M$
	$y^{mw}_{(od)}$	m 型公交车辆 w 为乘客 od 行程提供服务取 1,否则取 0。$(od) \in \psi, w \in W, m \in M$
	z^{mw}_i	m 型公交车辆 w 服务乘降点 i 时取 1,否则取 0。$w \in W, m \in M, i \in V$
	$\lambda[mw]$	表示 m 型公交车辆 w 线路需要访问的乘降点集合,是下层模型的优化目标。$w \in W, m \in M$
	α^{mw}_i	m 型公交车辆 w 的运营线路以乘降点 i 为始发站点时取 1,否则取 0。$i \in \lambda[od]^{mw}$
	β^{mw}_{ij}	m 型公交车辆 w 从乘降点 i 前往乘降点 j 时取 1,否则取 0。$i,j \in \lambda[od]^{mw}$
	γ^{mw}_i	m 型公交车辆 w 的运营线路以乘降点 i 为终到站点时取 1,否则取 0。$i \in \lambda[od]^{mw}$
	u^{mw}_i	m 型公交车辆 w 的运营线路中乘降点 i 的访问顺序
	$\sigma^{mw}_{(od)}$	m 型公交车辆完成 od 行程所需的时间,为整数决策变量。$(od) \in \psi, w \in W, m \in M$

1.3 双层混合整数规划模型构建

上层模型以考虑乘客出行成本和公交运营成本最小为优化目标。

$$\min \sum_{m\in M}\sum_{w\in W}\left(\xi_1^m \cdot f(g) + \sum_{\substack{i,j\in\lambda[mw]\\ j\neq i}} \xi_2^m \cdot \beta_{ij}^{mw} \cdot p_{ij} + \sum_{(od)\in\psi} c \cdot \sigma^{mw}_{(od)}\right) \tag{1}$$

$$\sum_{i,j\in\lambda[mw]} (\alpha_{ij}^{mw} \cdot p_{ij}) \leqslant l, \forall w \in W, \forall m \in M \tag{2}$$

$$\sum_{i\in\lambda[mw]} \alpha_i^{mw} = 1, \forall i \in \lambda[mw] \tag{3}$$

$$\sum_{i\in\lambda[mw]} \gamma_i^{mw} = 1, \forall i \in \lambda[mw] \tag{4}$$

$$\alpha_i^{mw} + \gamma_i^{mw} \leqslant 1, \forall i \in \lambda[mw] \tag{5}$$

$$\sum_{\substack{j=1,j\neq i\\ j\in\lambda[mw]}} \beta_{ji}^{mw} = 1 - \alpha_i^{mw}, \forall i \in \lambda[mw] \tag{6}$$

$$\sum_{\substack{j=1,j\neq i\\ j\in\lambda[mw]}} \beta_{ij}^{mw} = 1 - \gamma_i^{mw}, \forall i \in \lambda[mw] \tag{7}$$

$$u_j^{mw} \geqslant u_i^{mw} + 1 - n(1-\beta_{ij}^{mw}), \forall i \in \lambda[mw], i \neq j \tag{8}$$

$$u_i^{mw} \geqslant 1, u_i^{mw} \leqslant \| \lambda[mw] \| \tag{9}$$

式(1)为上层模型优化目标,其中第 1 部分是公交车固定发车成本,$f(g)$是公交车发车数量,为下层模型的优化目标;第 2 部分是公交车运行成本;第 3 部分是乘客出行的时间成本。式(2)为定制公交线路长度约束,要求定制公交车辆满足续航要求。式(3)~(5)为始发和终到站点唯一性约束。对任意定制公交,其运营线路只能有一个始发站点和终到站点,且同一乘降点不能同时作为始发站点和终到站点。式(6)为节点访问约束,对于定制公交的任意乘降点(非始发站点),都有定制公交车辆访问。式(7)同理,为节点离开约束,对于定制公交的任意乘降点(非终到站点),都有定制公交车辆离开。式(8)、(9)为子环路消除约束,$\| \lambda[mw] \|$是 m 型公交车辆 w 投入运营需要访问的乘降点数量。

$$\text{Min } g = \sum_{m\in M}\sum_{w\in W} x^{mw} \tag{10}$$

$$\sum_{m\in M}\sum_{w\in W} y^{mw}_{(od)} = 1, \forall (o,d) \in \psi \tag{11}$$

$$\sum_{(od)\in\psi} 2y^{mw}_{(od)} \leqslant \varphi, \forall w \in W, m \in M \tag{12}$$

$$\sum_{(o,d)\in\psi} y^{mw}_{(od)} \cdot n_{od} \geqslant \mu f_m, \forall w \in W, m \in M \tag{13}$$

$$\sum_{(o,d)\in\psi} y^{mw}_{(od)} \cdot n_{od} \leqslant f_m, \forall w \in W, m \in M \tag{14}$$

$$x^{mw} = y^{mw}_{(od)}, \forall w \in W, m \in M \tag{15}$$

$$z_o^{mw} = y^{mw}_{(od)}, z_d^{mw} = y^{mw}_{(od)}, \forall w \in W, m \in M, o,d \in V \tag{16}$$

$$\lambda[mw] = V \cdot z_i^{mw} \tag{17}$$

式(10)为下层模型的优化目标,是最小化公交车辆使用数量。式(11)为乘客 od 行程服务约束,要求任意乘客 od 行程都应该被一辆定制公交车服务。式(12)为线路乘降点访问数量约束,要求单条定制公交线路的乘降点访问数量不能超过最大访问数量限制。式(13)为最低乘车人数约束,要求公交线路的开通必须满足定制公交最低乘车人数限制。式(14)为最大运力约束,定制公交服务的乘客总数量不超过最大运力限制。式(15)为公交车投入运营约束,当 m 型公交车辆 w 服务乘客(od)行程时,该辆公交车应投入运营,否则无需投入运营。式(16)为乘降点访问约束,od 行程如果被定制公交车辆服务,该 od 行程的上车点 o 和下车点 d 应被纳入定制公交运营线路。式(17)定制公交线路服务站点集合生成。

2 基于 Actor-Critic 网络的深度强化学习

作为连接值函数逼近与策略优化两大技术路线的桥梁,基于 Actor-Critic 网络的深度强化学习近年来在动态调度、机器人控制等复杂决策领域的应用取得突破性进展。Actor-Critic 双重网络的深度强化学习算法将策略梯度方法与价值函数评估进行有机融合,能够直接优化策略并利用价值函数提供实时

反馈。通过 Critic 神经网络的价值评估，该算法能够在每一步决策后更新策略，这种单步更新机制显著提升了学习效率。Critic 神经网络的价值估计能够有效降低策略梯度估计中的方差，使训练过程更加稳定。此外，Critic 神经网络对策略潜在价值的评估能够指导探索过程，有助于解决强化学习中智能体的探索与利用平衡问题。与传统的 Q-learning 或策略梯度方法相比，Actor-Critic 双重网络的深度强化学习算法兼具高效性与稳定性，能够更好地适应复杂环境中的连续决策任务。

此外，多对多定制公交线网优化问题属于大规模车辆路径问题，精确求解效率不高。深度强化学习是深度学习和强化学习的结合，利用深度学习技术来处理高维复杂数据，并在强化学习框架下指导智能体在复杂环境中作出智能决策。本文设计基于 Actor-Critic 双重网络的深度强化学习方法。设置基于乘客数–乘降点–行驶距离的状态空间，构建基于接续乘降点选择的动作空间，给出基于乘客出行与公交运营的奖励函数，利用 Actor 神经网络（actor network，简称 AN）进行定制公交接续乘降点动作选择、Critic 神经网络（critic network，简称 CN）进行接续乘降点动作选择的优劣评价，给出基于乘降点选择与评价的 Actor-Critic 网络训练方法，形成基于 Actor-Critic 网络的深度强化学习过程。

2.1　定制公交智能体学习环境

以定制公交为智能体，乘降点位置和距离为定制公交智能体学习环境，使得定制公交智能体在学习环境的反馈奖励中不断优化策略以获得更大奖励值的动作优化策略。利用无监督聚类，依据上车点和下车点坐标位置，确定定制公交线路乘降点，构建公交智能体学习环境，步骤如下：

步骤 1　确定上车点坐标位置。通过定制公交乘客出行需求调查，收集乘客的经纬度位置信息，利用坐标转换工具将经纬度位置信息映射到平面坐标系。

步骤 2　生成上车点初始聚类划分。给定乘客步行距离，利用层次聚类法对上车点坐标位置数据进行聚类，得到 K 个上车点初始聚类划分。

步骤 3　更新上车点聚类划分。依次从 K 个初始聚类划分中随机选择 K 个上车点作为聚类中心，分别计算其余上车点与 K 个聚类中心之间的距离，并将上车点重新分配到距离最近的聚类中心所在划分，从而形成 K 个新聚类划分。

步骤 4　更新聚类划分中心上车点。对 K 个新聚类划分，分别计算聚类划分内所有上车点横坐标的平均值和纵坐标的平均值，并将横坐标的平均值和纵坐标的平均值作为新的聚类划分中心上车点。

步骤 5　上车点位置确定。重复执行步骤 3 和步骤 4，直至聚类划分中心位置不再改变，从而得到上车区域上车点位置。

步骤 6　下车区域下车点位置确定。对下车区域下车点坐标位置数据执行上述步骤，从而得到定制公交线路的下车区域下车点位置。

定制公交智能体与学习环境交互可用马尔科夫决策过程表示，即下一状态取决于智能体当前状态和采取的动作。以定制公交智能体从当前乘降点到接续乘降点为时段 t，下面对定制公交智能体的状态–动作–奖励机制进行设置。

2.2　定制公交智能体状态–动作–奖励机制

2.2.1　基于乘客数–乘降点–距离的状态空间

设定定制公交的乘客数、所在乘降点和已行驶距离为状态空间 S_t，以表示时段 t 定制公交智能体与学习环境进行交互产生的信息，即

$$S_t = \{ n_t^{mw}, v_{\text{start}}^{mw}, v_{\text{end}}^{mw}, v_\theta^{mwt}, b_t^{mw} \}$$

式中：n_t^{mw} 表示 m 型定制公交 w 在时段 t 的乘客数，v_{start}^{mw} 表示 m 型定制公交 w 运营线路的始发站点；v_{end}^{mw} 表示 m 型定制公交 w 运营线路的终到站点；v_θ^{mwt} 表示时段 t 定制公交智能体所处状态，当 $\theta = 0$ 时，表示智能体处在运营线路的始发站点，当 $\theta = 1$ 时，表示智能体处在两个乘降点途中，当 $\theta = 2$ 时，表示智能体处在运营线路的终到站点；b_t^{mw} 表示 m 型定制公交 w 在时段 t 行驶距离。

2.2.2　基于接续乘降点选择的动作空间

设定定制公交智能体状态为 S_t 时，即乘客数为 n_t^{mw}、乘降点状态为 v_θ^{mwt} 和已行驶距离为 b_t^{mw} 时，可选择的下一接续乘降点集合为动作空间 A_t，即

$$A_t = \{ v \mid v \in V \}$$

式中 v 表示定制公交智能体可选择下一乘降点的动作。

2.2.3　基于乘客出行与公交运营的奖励函数

设定定制公交乘客出行成本 Z_1 的倒数与公交运营成本 Z_2 的倒数之和为奖励函数 R_t，以表示时段 t 定制公交智能体与学习环境交互后，执行动作后获得奖励值，即 $R_t = 1/Z_1 + 1/Z_2$。

2.3　基于乘降点选择与评价的 Actor-Critic 网络训练

利用 Actor 网络（actor network，简称 AN）进行定制公交接续乘降点动作选择，Critic 网络（critic network，简称 CN）进行选择不同接续乘降点的概率分布评价，如图 1 所示，具体步骤如下：

步骤 1　Actor-Critic 网络训练的超参数设置。设置学习率 α、折扣因子 γ、批尺寸 β、神经元数 η 作为 Actor-Critic 网络训练的超参数。

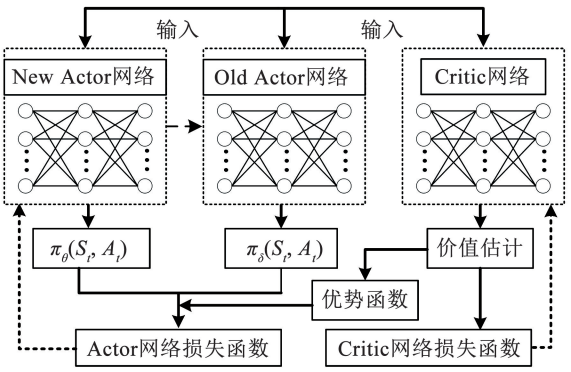


图 1 Actor-Critic 网络训练过程

Fig. 1 Training process of the Actor-Critic network

步骤 2 基于接续乘降点选择的 Actor 网络输入输出设置。Actor 网络分为用于更新控制参数的 New Actor 网络和用于收集经验数据的 Old Actor 网络,设置其控制参数分别为 θ 和 δ,将定制公交状态空间 S_t 和动作空间 A_t 作为 New Actor 网络和 Old Actor 网络输入,输出定制公交从动作空间 A_t 中选择不同接续乘降点的概率分布 $\pi_\theta(S_t,A_t)$ 和 $\pi_\delta(S_t,A_t)$。

步骤 3 基于评价不同接续乘降点概率分布的 Critic 网络输入输出设置。设置 Critic 网络的控制参数 λ,将 Old Actor 网络的状态空间 S_t 和动作空间 A_t 作为 Critic 网络输入,输出状态空间 S_t 的价值估计值 $H(S_t|\pi_\delta)$ 和动作空间 A_t 的价值估计值 $U(A_t|\pi_\delta)$。

步骤 4 Critic 神经网络训练。利用优势函数 $\psi(S_t,A_t)$ 评估在状态 S_t 下执行选择不同乘降点的动作 A_t 对于当前策略的价值提升,如式(18)所示。采用时序差分误差的方法计算 Critic 网络损失函数 L^{CN},并反向传播更新参数 λ,如式(19)所示。通过最小化损失函数,Critic 神经网络更准确地估计状态价值,从而为 Actor 网络提供反馈,即

$$\psi(S_t,A_t)=U(A_t|\pi_\delta)-H(S_t|\pi_\delta) \tag{18}$$

$$L^{CN}(\lambda)=E_t(R_t+\gamma H(S_{t+1}|\pi_\delta)-H(S_t|\pi_\delta))^2 \tag{19}$$

式中 $E_t(\ )$ 为 t 时段的期望函数;γ 为折扣因子;$H(S_{t+1}|\pi_\delta)$ 为下一状态 S_{t+1} 的价值估计值;$\psi(S_t,A_t)$ 为定制公交状态 S_t 下采取动作 A_t 的优势函数,当 $\psi(S_t,A_t)>0$ 时,说明此动作优于平均动作,增加选择该动作的概率;当 $\psi(S_t,A_t)<0$ 时,说明此动作劣于平均动作,减少选择该动作的概率。

步骤 5 Actor 神经网络训练。计算 New Actor 和 Old Actor 网络概率分布比率 τ 和 Actor 网络损失函数 L^{AN} 并反向传播更新 New Actor 网络参数 θ,间隔固定迭代次数 ς 复制 New Actor 网络参数 θ 来更新 Old Actor 网络参数 δ,即

$$\tau=\frac{\pi_\vartheta(S_t,A_t)}{\pi_\delta(S_t,A_t)} \tag{20}$$

$$L^{AN}(\theta)=-E_t[\min(\tau\cdot\psi(S_t,A_t),\psi(S_t,A_t)\cdot\mathrm{clip}(\tau,1-\varepsilon,1+\varepsilon))] \tag{21}$$

式中 min(·)为最小值函数;τ 为 New Actor 和 Old Actor 网络概率分布比率;clip 为 Python 内置截断函数,将 New Actor 和 Old Actor 网络参数更新限定在 $[1-\varepsilon,1+\varepsilon]$ 之间;ε 为截断系数,参照张玺君等[19]将该系数设置为 0.2。

2.4 基于 Actor-Critic 网络的深度强化学习过程

基于 Actor-Critic 网络深度强化学习的定制公交线网优化过程,如图 2 所示,具体步骤如下:

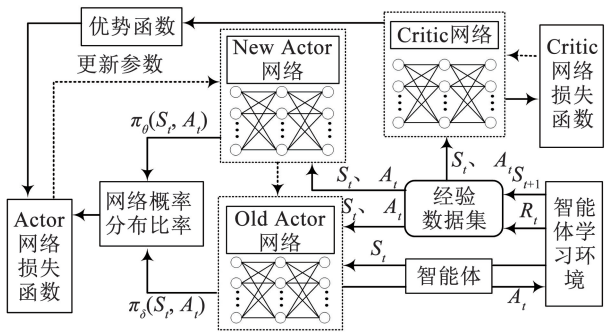


图 2 基于 Actor-Critic 网络的深度强化学习过程

Fig. 2 Deep reinforcement learning process based on the Actor-Critic network

步骤 1 初始化定制公交智能体学习环境、基于乘客数-乘降点-行驶距离的状态空间和基于接续乘降点选择的动作空间。

步骤 2 设置用于 Actor-Critic 网络训练的超参数,即学习率 α、折扣因子 γ、批尺寸 β、神经元数 η。

步骤 3 随机初始化生成 New Actor 网络控制参数 θ、Old Actor 网络控制参数 δ 和 Critic 网络控制参数 λ。

步骤 4 根据定制公交状态空间 S_t,智能体选择不同接续乘降点的动作空间 A_t,得到奖励 R_t 和下一时段状态空间 S_{t+1}。

步骤 5 记录每一时段 t 的当前状态空间 S_t、下一状态空间 S_{t+1}、动作空间 A_t 以及奖励 R_t 构造并存放到经验数据集。

步骤 6 从经验数据集中随机抽取 β 个数据用于训练 Actor-Critic 网络训练。

步骤 7 利用式(1)计算评价定制公交智能体接续乘降点动作好坏的优势函数 $\psi(S_t,A_t)$。

步骤 8 利用式(2)计算 Critic 网络损失函数值,使用梯度下降法更新参数 λ,以最小化损失函数。

步骤 9 利用式(3)计算 New Actor 和 Old Actor 网络概率分布比率 τ。

步骤 10 利用式(4)计算 Actor 网络损失函数值,反向传播更新 New Actor 网络控制参数 θ,间隔

固定迭代次数 s 复制 New Actor 网络控制参数 θ 更新 Old Actor 网络参数 δ。

步骤 11　判断迭代次数是否达到最大迭代次数，若未达到，则转到步骤 4，直至满足最大迭代次数结束。

3　案例研究与结果分析

3.1　数据采集与处理

借助公共交通大数据平台，并通过实地调研郑州市高新区、郑东新区乘客流量相对集中的商圈、居住小区和工作区出行数据，同时利用百度地图捕获乘客出行起讫点经纬度坐标，借助 COORD 软件，利用高斯三度带地图投影方法，将乘客出行经纬度坐标转换为平面坐标，其分布如图 3 所示。

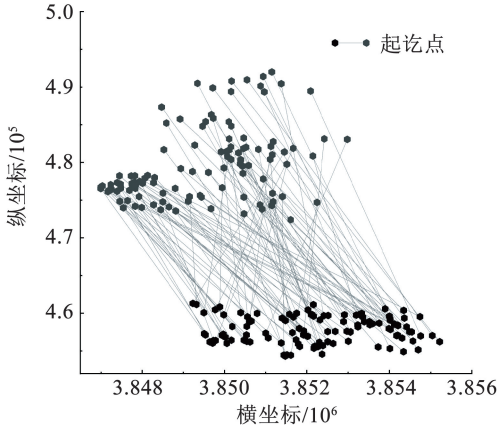


图 3　定制公交乘客起讫点转换平面坐标分布

Fig. 3　Distribution of converted planar coordinates of customized bus passenger origins and destinations

定制公交车队的车辆配属规模为 6 辆，大小 2 种车型配属，小车型满载 18 人，大车型满载 30 人，单次发车成本分别为 80 元/车和 100 元/车，单位距离运行成本为 2 元/km，行驶速度为 35 km/h，单条线路乘降点最大允许访问数量限制为 6 个，公交车辆最低上座率为 75%，单条线路最长运行距离为 40 km，乘客在车时间成本为 6 元/(人·h)。

3.2　基于无监督聚类的智能体学习环境构建

利用层次聚类法对乘客出行起讫点坐标位置数据进行聚类，给定乘客步行距离 800 m，得到 8 个上车点初始聚类划分和 6 个下车点初始聚类划分。依次更新上车点和下车点聚类划分及其聚类划分中心，得到上车点聚类划分中心位置和下车点聚类划分中心位置，如图 4、5 所示。

对上下车区域聚类划分结果进行评估，计算出上车区域聚类平均轮廓系数为 0.451，下车区域聚类平均轮廓系数为 0.500，达到较好聚类效果。通过无监督聚类分析获取乘降点的平面坐标，利用 COORD 坐标转换工具将平面坐标转换为经纬度坐标，依据上下车点附近常规公交站点和道路环境，调整站点的位置，得到满足定制公交乘客出行要求的上下车点以及距离矩阵，如表 3 和表 5 所示，其中编号为 o 开头的表示上车点，标号以 d 开头的表示下车点。

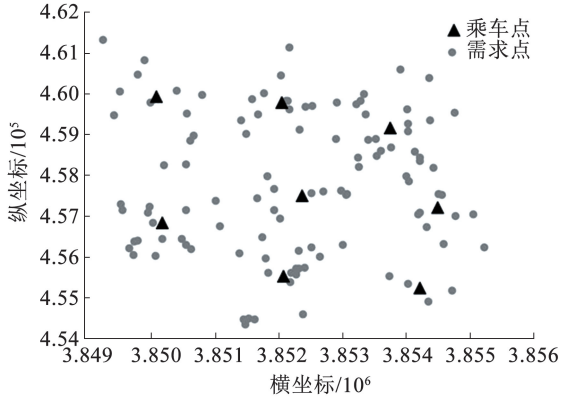


图 4　上车点聚类划分中心位置转换坐标分布

Fig. 4　Distribution of converted coordinates of boarding-stop cluster centers

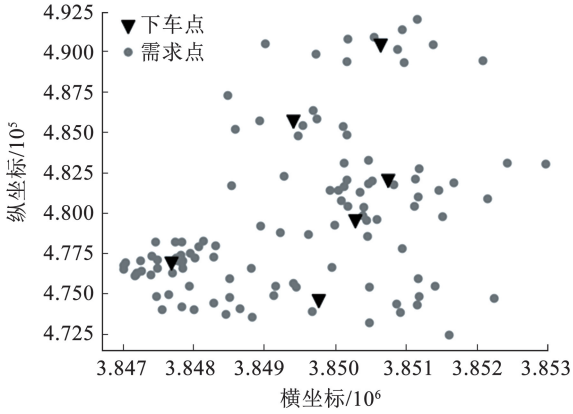


图 5　下车点聚类划分中心位置转换坐标分布

Fig. 5　Distribution of converted coordinates of alighting-stop cluster centers

表 3　定制公交上下车站点的经纬度

Tab. 3　Longitude and latitude of customized bus boarding and alighting stops

编号	乘降点	经纬度/(°)
*o*1	化工路	113.562 331,34.777 571
*o*2	民生纺织	113.528 301,34.778 251
*o*3	金色嘉园	113.560 498,34.795 360
*o*4	长椿路合欢街	113.535 516,34.798 129
*o*5	紫竹路梧桐街	113.513 859,34.795 341
*o*6	保利文化广场	113.553 545,34.810 629
*o*7	长椿路翠竹路	113.532 240,34.817 244
*o*8	紫兰苑	113.510 604,34.814 610
*d*9	商鼎路七里河南路	113.748 567,34.756 602
*d*10	中建七局	113.843 176,34.772 286
*d*11	如意湖	113.722 063,34.773 719
*d*12	东三环	113.775 845,34.779 970
*d*13	智慧城云著	113.907 112,34.788 412
*d*14	明理路龙子湖南路	113.815 883,34.788 974

进一步的,将 8 个上车点依次执行以下步骤:首先,统计上车点覆盖范围内所有乘客 OD 行程;接着,将相同目的地(即相同下车点)的 OD 行程合并,得到合并后的 OD 行程及相应客流量,如表 4 所示。

3.3 Actor-Critic 网络训练的超参数设置

利用 Anaconda 平台使用 python 在 Intel(R) Core(TM)i5-8250U CPU@1.60 GHz 1.80 GHz 环境下编译基于 Actor-Critic 网络的深度强化学习算法,对定制公交智能体进行训练学习。设定最大迭代次数为 100 000,利用 New Actor 网络控制参数更新 Old Actor 网络控制参数的间隔固定迭代次数 $\varsigma=1\ 000$。

为实现基于 Actor-Critic 网络的深度强化学习算法的最佳性能,对学习率 α、折扣因子 γ、批尺寸 β 和神经元数 η 四个超参数进行对比验证,分析超参数对训练速率的影响,如图 6 所示。横轴表示迭代次数,纵轴表示训练速率。

表 4 乘客出行 OD 行程信息

Tab. 4 Origin-destination (OD) trip information of passengers

行程编号	OD 行程	出行人数
1	($o1$,$d14$)	3
2	($o2$,$d9$)	14
3	($o2$,$d11$)	8
4	($o3$,$d9$)	5
5	($o4$,$d10$)	5
6	($o4$,$d14$)	10
7	($o5$,$d10$)	3
8	($o5$,$d11$)	12
9	($o5$,$d12$)	10
10	($o5$,$d14$)	4
11	($o6$,$d9$)	10
12	($o6$,$d12$)	6
13	($o7$,$d13$)	8
14	($o8$,$d11$)	6
15	($o8$,$d13$)	7

表 5 上下车站点距离矩阵

Tab. 5 Distance matrix of boarding and alighting stops

乘降点	上下车站点距离/km													
	$o1$	$o2$	$o3$	$o4$	$o5$	$o6$	$o7$	$o8$	$d9$	$d10$	$d11$	$d12$	$d13$	$d14$
$o1$	0	3.1	1.9	3.3	4.8	3.7	5.1	6.2	17.1	25.6	14.6	19.5	31.5	23.2
$o2$	3.1	0	3.5	2.3	2.4	4.2	4.4	4.5	20.3	28.8	17.7	22.6	34.6	26.3
$o3$	1.9	3.5	0	2.3	4.2	1.8	3.5	5.0	17.7	25.9	14.9	19.7	31.6	23.3
$o4$	3.3	2.3	2.3	0	2.0	2.2	2.1	2.9	19.9	28.2	17.2	22.0	33.9	25.6
$o5$	4.8	2.4	4.2	2.0	0	4.0	2.9	2.2	21.8	30.1	19.1	23.9	35.9	27.5
$o6$	3.7	4.2	1.8	2.2	4.0	0	2.0	3.9	18.8	26.7	15.9	20.5	32.3	24.0
$o7$	5.1	4.4	3.5	2.1	2.9	2.0	0	1.9	20.8	28.8	17.9	22.6	34.3	26.0
$o8$	6.2	4.5	5.0	2.9	2.2	3.9	1.9	0	22.6	30.7	19.8	24.5	36.3	28.0
$d9$	17.1	20.3	17.7	19.9	21.8	18.8	20.8	22.6	0	8.8	3.0	3.6	14.9	7.1
$d10$	25.6	28.8	25.9	28.2	30.1	26.7	28.8	30.7	8.8	0	11.0	6.2	6.1	3.1
$d11$	14.6	17.7	14.9	17.2	19.1	15.9	17.9	19.8	3.0	11.0	0	4.9	16.9	8.7
$d12$	19.5	22.6	19.7	22.0	23.9	20.5	22.6	24.5	3.6	6.2	4.9	0	12.0	3.7
$d13$	31.5	34.6	31.6	33.9	35.9	32.3	34.3	36.3	14.9	6.1	16.9	12.0	0	8.3
$d14$	23.2	26.3	23.3	25.6	27.5	24.0	26.0	28.0	7.1	3.1	8.7	3.7	8.3	0

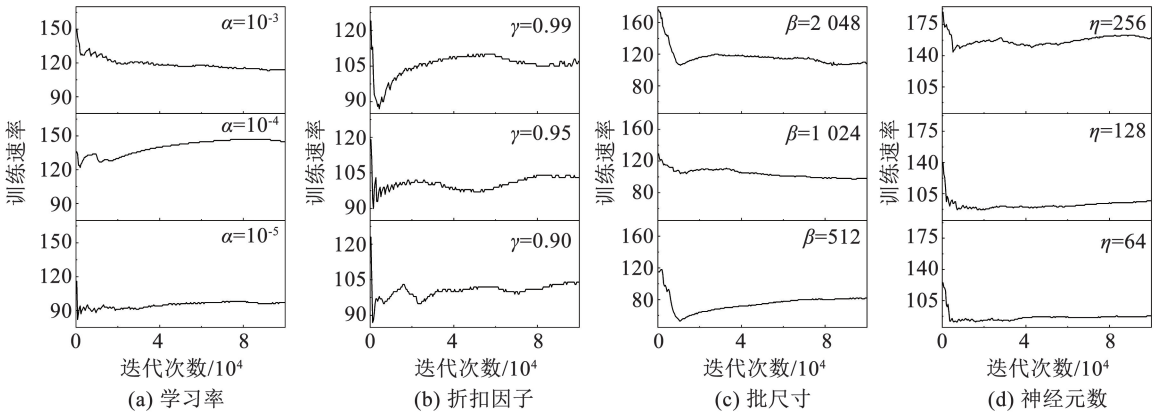


图 6 超参数不同取值对训练速率的影响

Fig. 6 Effects of different hyperparameter settings on training speed

图 6(a)表示学习率不同取值对训练速率的影响。当学习率 $\alpha=1\times10^{-3}$时，训练速率值呈现下降趋势；当学习率 $\alpha=1\times10^{-5}$时，训练速率值在 80～100 低值之间波动；当学习率 $\alpha=1\times10^{-4}$时，训练速率值呈现上升趋势，且数值明显高于学习率为 1×10^{-3}、1×10^{-5}。故学习率选取 1×10^{-4}。

图 6(b)表示折扣因子不同取值对训练速率的影响。当折扣因子 $\gamma=0.90$ 时，训练速率值波动于 85～105 的低值之间；当折扣因子 $\gamma=0.95$ 时，训练速率值有所提升，但训练速率值仍较低；当折扣因子 $\gamma=0.99$ 时，训练速率值呈现先下降后回升的趋势同时保持 105～110 的高值之间。故折扣因子选取 0.99。

图 6(c)表示批尺寸不同取值对训练速率的影响。当批尺寸 $\beta=512$ 时，训练速率值呈现大幅度下降后小幅度回升，其数值稳定在 60～80 之间；当批尺寸 $\beta=1\ 024$ 时，训练速率值出现显著的下降趋势；当批尺寸 $\beta=2\ 048$ 时，训练速率值呈现先下降后回升的趋势，且数值均高于批尺寸为 512、1 024。故批尺寸选取 2 048。

图 6(d)表示神经元数不同取值对训练速率的影响。当神经元数 $\eta=64$、128 时，训练速率值均先下降后稳定在一定的区间内；当神经元数 $\eta=256$ 时，训练速率值呈现微小波动，且数值都高于学习率为 64、128 时的表现。故神经元数选取 256。

通过对学习率 α、折扣因子 γ、批尺寸 β、神经元数 η 进行灵敏度分析，超参数取值如表 6 所示。

表 6　超参数取值

Tab. 6　Hyperparameter settings

超参数	取值
学习率 α	1×10^{-4}
折扣因子 γ	0.99
批尺寸 β	2 048
神经元数 η	256

3.4　研究结果分析

利用基于 Actor-Critic 网络的深度强化学习进行求解，图 7 表示迭代次数与奖励值的收敛效果图，迭代次数达到 2×10^4 后奖励值收敛且平稳保持在 0.003 2 附近，寻优能力强且收敛速度快。当训练次数大于 2×10^4 次后，Actor 网络的损失值逐渐趋于 0，Critic 网络的损失函数值在 0.016～0.025 之间震荡，说明基于 Actor-Critic 网络深度强化学习的公交智能体训练效果较好。

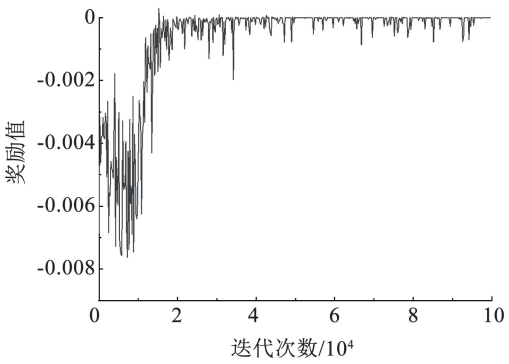


图 7　迭代次数与奖励值的收敛效果

Fig. 7　Convergence of reward values over iterations

如图 7 所示，奖励值在迭代次数 2×10^4 后趋于稳定，计算公交运营成本 $Z_1=756.4$ 元，乘客出行成本 $Z_2=506.3$ 元，得到定制公交线网优化方案如表 7 和定制公交线路分布如图 8 所示。

表 7　定制公交线网优化方案

Tab. 7　Optimization scheme of the customized bus network

编号	运营线路	行驶距离/km	乘降点数量	上座率/%	车型(载客量)
1	8—7—13	36.2	3	83.3	18
2	5—4—6—9—10	31.8	5	100	18
3	2—5—8—11	24.4	4	86.7	30
4	5—4—1—12—14	28.5	5	90.0	30
5	2—6—3—9—12	27.3	5	83.3	30

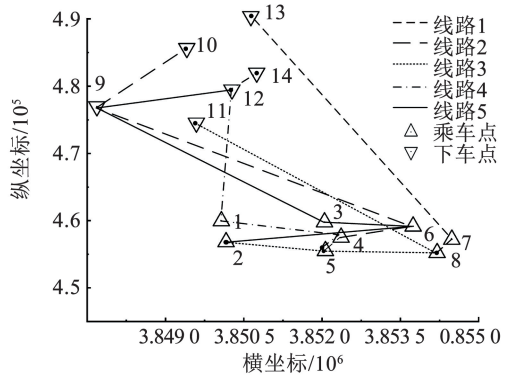


图 8　定制公交线路分布

Fig. 8　Spatial distribution of customized bus routes

由表 7 可以看出，定制公交线网优化方案的线路平均上座率为 88.66%，共配属 2 辆载客量为 18 的公交和 3 辆载客量为 30 的公交，车辆行驶距离均控制在合理范围内，且线路中乘降点个数均未超过 6 个，从而保障了定制公交一人一座、舒适直达的优势。

从定制公交运行时间来看，图 9 表示定制公交不同线路的运行时间分布图，任意一辆定制公交完成服务所需最长时间均小于 66 min，5 辆车的总运行时间为 276 min。

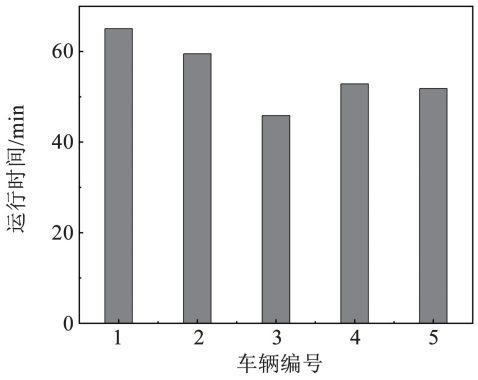


图 9 定制公交运行时间分布

Fig. 9 Distribution of customized bus travel times

从定制公交车辆上座率来看,图 10 表示定制公交不同线路的上座率分布图,5 辆定制公交的上座率均超过 80%,其中有 1 辆定制公交车的上座率接近或达到 100%,充分利用车辆资源。

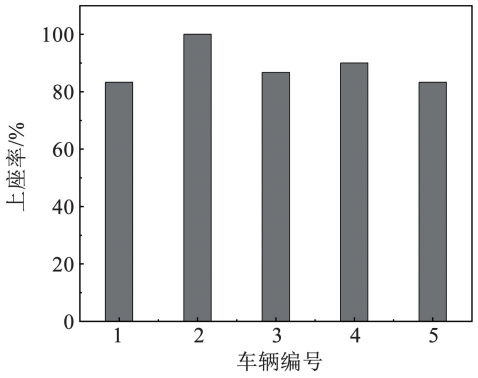


图 10 定制公交车辆上座率分布

Fig. 10 Distribution of customized bus vehicle load factors

通过分析与对比,利用 Actor-Critic 双重网络深度强化学习算法得到的定制公交开行方案,其车辆行驶时长、车辆上座率与车型安排等指标数据均反映出较好的性能。

3.5 Actor-Critic 网络与 Actor 网络对比

为了验证基于 Actor-Critic 双重网络深度强化学习的可行性和优越性,将其与基于 Actor 单重网络的深度强化学习进行对比,采取相同的状态、动作、奖励机制和相同的超参数取值进行训练学习,对比结果如表 8 所示,不同算法下迭代次数与奖励值的关系如图 11 所示。由表 8 可知,基于 Actor 单重网络的深度强化学习的计算时间为 6 h,而基于 Actor-Critic 双重网络深度强化学习的计算时间为 4.5 h,相比于 Actor 单重网络的深度强化学习减少了 1.5 h。

较基于 Actor 单重网络的深度强化学习算法,基于 Actor-Critic 双重网络的深度强化学习算法收敛速度快且奖励值高,其原因在于 Critic 网络具有提高算法稳定性、加速收敛和提高样本效率的优势,提高定制公交智能体学习的效率和稳定性。

表 8 Actor-Critic 网络与 Actor 网络对比结果

Tab. 8 Performance comparison between the Actor-Critic network and the Actor network

类型	运行时间/h	奖励值迭代次数	损失值迭代次数
Actor 网络	6.0	70 000	70 000
Actor-Critic 网络	4.5	20 000	50 000

图 11 表示不同算法下迭代次数与奖励值的对比关系,基于 Actor-Critic 双重网络深度强化学习算法的奖励值在迭代次数为 0 ~ 20 000 之间时波动性较大,但在 20 000 ~ 50 000 之间时波动性逐渐降低,50 000 次之后趋于稳定;而基于 Actor 单重网络深度强化学习算法的波动性在迭代次数为 0 ~ 20 000 之间没有趋于稳定,直至 70 000 次之后趋于稳定收敛。基于 Actor-Critic 双重网络深度强化学习算法较基于 Actor 单重网络深度强化学习算法的收敛速度提高了 28.57%,奖励值增加了 13.65%,说明本文算法能够快速高效解决多对多定制公交线网优化问题。

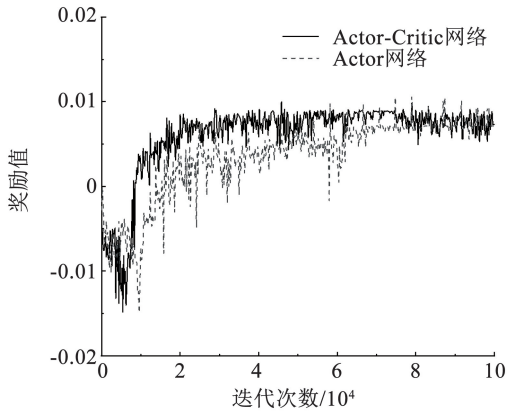


图 11 不同算法下迭代次数与奖励值的关系

Fig. 11 Reward convergence under different algorithms

3.6 Actor-Critic 网络强化学习算法求解性能评估

为进一步验证 Actor-Critic(为便于表述,简写为 AC)网络强化学习算法的高效性,本节引入遗传算法(genetic algorithm, GA)和邻域搜索算法(Neighborhood Search,NS)作为对比,并导入五种更大规模试验中进行测试。同时选取行驶距离、总运营成本和运行时间作为指标对不同算法的求解性能进行试验比对。GA 和 NS 为十次求解的平均值。这里改进率 η_1 =(GA 算法的总运营成本 - AC 的总运营成本)/AC 的总运营成本 ×100%,改进率 η_2 =(NS 算法的总运营成本 - AC 的总运营成本)/AC 的总运营成本 ×100%。5 种不同规模试验场景下的连续 10 次测试得到的平均值如表 9 所示。

表 9　5 种不同规模试验场景下的算法性能比对

Tab. 9　Performance comparison of algorithms across five test scenarios of different scales

算例	出行人数	实际用车数		总运营成本/元			改进率	
		小车型	大车型	AC	GA	NS	η_1	η_2
11O-10D	122	3	3	1 481	1 558	1 537	5.2	3.8
13O-15D	154	3	4	1 875	2 010	1 981	7.2	5.7
15O-17D	183	5	4	2 303	2 536	2 492	10.1	8.2
20O-18D	190	6	4	2 491	2 780	2 733	11.6	9.7
25O-22D	227	5	6	2 883	3 248	3 185	12.7	10.5

由表 9 中数据可以看出：1）由总运营成本指标可知，AC 求解的总运营成本最低、NS 次之、GA 最高。总运营成本指标比较结果为 AC < NS < GA。2）由改进率指标可知，小规模场景下，AC 相较 GA 和 NS 的求解优越性不突出；但随着规模逐渐增加，AC 的优越性愈发明显。

综上，Actor-Critic 双重网络深度强化学习算法在模型求解质量上具有明显优势，尤其随着问题规模的增加，优势愈加明显。

4　结　论

本文研究多对多定制公交线网优化问题，以乘客出行成本和公交运营成本最小为目标，考虑乘降点服务、乘降点数量、线路总长度以及公交车队规模等限制，构建多对多定制公交线网优化模型。

鉴于多对多定制公交线网优化问题属于大规模车辆路径问题，精确求解效率不高。故提出基于 Actor-Critic 双重网络的深度强化学习方法。该方法将定制公交视为智能体，利用无监督聚类建立智能体学习环境，构造智能体状态–动作–奖励机制，设计基于乘降点选择与评价的 Actor-Critic 神经网络，搭建基于 Actor-Critic 网络深度强化学习过程。通过选取郑州市高新区–郑东新区为实验场景，对学习率、折扣因子、批尺寸和神经元数等超参数进行优化调试，求解得到多对多定制公交线网优化方案，并通过与其他算法对比，得到如下结论：

1）与基于 Actor 单重网络的深度强化学习进行对比，算法收敛速度提高了 28.57%，奖励值增加了 13.65%，从而更加快速高效解决多对多定制公交线网优化问题。

2）与启发式算法对比结果表明，本文所提出的方法可以快速高效完成多对多定制公交线网优化模型求解，合理确定定制公交乘降点位置和具体线路走向，最大限度解决乘客通勤需求问题，缩减乘客的出行时间成本和公交企业的运营成本。

考虑定制公交行程时间的不确定性以及乘客出行行程取消等因素扰动，对定制公交线网优化模型进一步完善，是未来进一步的研究方向。

参考文献

[1] Liu T, Ceder A. Analysis of a new public-transport-service concept: customized bus in China[J]. Transport Policy, 2015, 39: 63

[2] Wang J B, Yamamoto T, Liu K. Role of customized bus services in the transportation system: insight from actual performance [J]. Journal of Advanced Transportation, 2019, 2019: 6171532

[3] Wang J B, Yamamoto T, Liu K. Spatial dependence and spillover effects in customized bus demand: empirical evidence using spatial dynamic panel models[J]. Transport Policy, 2021, 105: 166

[4] Ellegood W A, Campbell J F, North J. Continuous approximation models for mixed load school bus routing [J]. Transportation Research Part B: Methodological, 2015, 77: 182

[5] 安久煜, 宋瑞, 毕明凯, 等. 高铁车站接驳公交灵活线路优化设计研究[J]. 交通运输系统工程与信息, 2019, 19(5): 150
An Jiuyu, Song Rui, Bi Mingkai, et al. Optimization flexible route design for high-speed railway station feeder bus [J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(5): 150

[6] 王正武, 向健, 喻杰. 响应型接驳公交系统基于关键点的动态路径优化[J]. 长沙理工大学学报(自然科学版), 2020, 17(3): 51
Wang Zhengwu, Xiang Jian, Yu Jie. Dynamic route optimization based on key points for responsive feeder transit system[J]. Journal of Changsha University of Science & Technology (Natural Science), 2020, 17(3): 51

[7] Guo R G, Guan W, Zhang W Y. Route design problem of customized buses: mixed integer programming model and case study [J]. Journal of Transportation Engineering, Part A: Systems, 2018, 144(11): 04018069

[8]马昌喜, 王超, 郝威,等. 突发公共卫生事件下应急定制公交线路优化[J]. 交通运输工程学报, 2020, 20(3): 89
Ma Changxi, Wang Chao, Hao Wei, et al. Emergency customized bus route optimization under public health emergencies[J]. Journal of Traffic and Transportation Engineering, 2020, 20(3): 89

[9]王超, 马昌喜. 基于遗传算法的定制公交多停车场多车线路优化[J]. 交通信息与安全, 2019, 37(3): 109
Wang Chao, Ma Changxi. Optimization of parking lot and multi-vehicle route of customized buses based on genetic algorithm[J]. Journal of Transport Information and Safety, 2019, 37(3): 109

[10]Li Z J, Song R, He S W, et al. Methodology of mixed load customized bus lines and adjustment based on time windows[J]. Plos One, 2018, 13(1):

[11]Wang C, Ma C X, Xu X C. Multi-objective optimization of real-time customized bus routes based on two-stage method[J]. Physica A: Statistical Mechanics and its Applications, 2020, 537: 122774

[12]陈汐, 王印海, 刘剑锋, 等. 多区域通勤定制公交线路规划模型及求解算法[J]. 交通运输系统工程与信息, 2020, 20(4): 166
Chen Xi, Wang Yinhai, Liu Jianfeng, et al. Multi-region commuting bus route design model and solution method[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(4): 166

[13]Huang K, Xu L, Chen Y, et al. Customized bus route optimization with the real-time data[J]. Journal of Advanced Transportation, 2020, 2020: 8838994

[14]胡郁葱, 陈栩, 罗嘉陵. 多起终点多车型混载的定制公交线路规划模型[J]. 广西师范大学学报(自然科学版), 201, 36(4): 1
Hu Yucong, Chen Xu, Luo Jialing. Network design model of customized bus in diversified operation of multi-origin-destination and multi-type vehicle mixed load[J]. Journal of Guangxi Normal University (Natural Science Edition), 2018, 36(4): 1

[15]何民, 李沐轩, 税文兵,等. 可靠性和舒适性对响应式定制公交线路设计的影响[J]. 公路交通科技, 2019, 36(5): 145
He Min, Li Muxuan, Shui Wenbing, et al. Influence of reliability and comfort on responsive custom bus route design[J]. Journal of Highway and Transportation Research and Development, 2019, 36(5): 145

[16]Wang A G, Guan H Z, Wang P F. Cross-regional customized bus route planning considering staggered commuting during the COVID-19[J]. IEEE Access, 2021, 9: 20208

[17]彭理群, 罗明波, 卢赫,等. 基于 Q-learning 的定制公交跨区域路径规划研究[J]. 交通运输系统工程与信息, 2020, 20(1): 104
Peng Liqun, Luo Mingbo, Lu He, et al. Cross-regional customized bus path planning based on Q-learning [J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(1): 104

[18]马昌喜, 郝威, 沈金星, 等. 定制公交线路优化综述[J]. 交通运输工程学报, 2021, 21(5): 30
Ma Changxi, Hao Wei, Shen Jinxing, et al. Review on customized bus route optimization[J]. Journal of Traffic and Transportation Engineering, 2021, 21(5): 30

[19]张玺君, 聂生元, 李喆, 等. 基于自注意力机制的深度强化学习交通信号控制[J]. 交通运输系统工程与信息, 2024, 24(2): 96
Zhang Xijun, Nie Shengyuan, Li Zhe, et al. Traffic signal control with deep reinforcement learning and self-attention mechanism[J]. Journal of Transportation Systems Engineering and Information Technology, 2024, 24(2): 96