

# 引力视角下出租车调度的增强学习模型

崔允汀, 何胜学\*

上海理工大学管理学院, 上海

收稿日期: 2021年9月19日; 录用日期: 2021年10月26日; 发布日期: 2021年11月2日

## 摘要

以互联网预约式叫车服务为背景, 针对出租车运营中司乘匹配率低、乘客等车时间长和车辆空驶距离长的问题, 通过类比万有引力模型, 提出了空驶出租车巡游路线调度优化的增强学习控制方法。分别以备选目的地等车乘客数和当前位置空驶出租车数为对应位置的“质量”, 而以两地间最短路径长度为两点间“距离”, 定义两点间“引力”。通过引力大小, 决策空驶车辆的目的地。通过引入增强学习理论, 实现空驶出租车巡游路线决策的时空全局最优性, 从而避免仅凭当前引力值大小决策路线的短视性。数值算例分析表明: 新的调度方法不仅可以提高司乘匹配率和总的运营收入, 而且可以减少总的空驶距离。

## 关键词

城市交通, 出租车调度, 增强学习, 引力模型

# Reinforcement Learning Model of Taxi Dispatching under Gravity Perspective

Yunting Cui, Shengxue He\*

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Sep. 19<sup>th</sup>, 2021; accepted: Oct. 26<sup>th</sup>, 2021; published: Nov. 2<sup>nd</sup>, 2021

## Abstract

In the setting of internet appointing taxi service, by analogizing the universal gravitation model, a control method of optimizing the cruising routes of vacant taxis was proposed to solve the problems of the low matching rate between taxis and passengers, the long average waiting time of passengers and the long travelling distance of vacant taxis. Using the number of waiting passengers as the mass of the corresponding alternative destination, the number of vacant taxis as the mass of

\*通讯作者。

current position, and the length of the shortest path between these two positions as their distance, the gravitation between these two positions can be defined. The next destination of a vacant taxi can be determined using the defined gravitation. By introducing the reinforcement learning theory, the decision of choosing the routes of vacant taxis possesses the global optimality. The short sight of determining the routes only by the size of gravitation can also be avoided. The numerical example shows that the new taxi dispatching method can not only increase the matching rate between taxis and passengers and the total operation income, but also decrease the total travelling distance of vacant taxis.

## Keywords

Urban Transportation, Taxi Dispatching, Reinforcement Learning, Gravity Model

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着手机叫车 APP 的普及, 过去的扬招式出租车服务逐渐被各种互联网预约式叫车服务所代替。GPS 和 GIS 技术也为出租车管理部门或运行公司提供了各种车辆运行实时信息和乘客需求信息。如何有效利用上述信息进一步提升出租车服务水平日益成为出租车运营公司谋求发展和重构核心竞争力的关键。如何利用现在信息通讯技术所带来的变化机遇实现运营服务各项关键指标的提升业已成为出租车运营企业面临的头号技术问题。在调度操作层面, 企业面临的核心问题就是如何通过实时动态的空驶车辆调度控制, 减少出租车空驶率和乘客的等车时间, 并增加司机匹配成功率和企业的运营收入。为了解决上述核心问题, 本文以当前出租车行业发展的现状为背景, 通过类比牛顿的万有引力模型, 提出了一种新的空车路线选择方法, 并结合增强学习理论实现对出租车服务系统的实时动态优化控制。

当一辆空驶的出租车选择下一时刻巡游目的地时, 会考虑到到达下一目的地后成功揽客的机率。成功揽客的机率取决于到达选定目的地时的等车乘客数量和当前位置与该目的地之间的距离。当然揽客成功与否也和同样选择该目的地的其他竞争出租车的数量相关。如果我们作如下的类比, 将上述等车乘客数比做目的地的“质量”, 当前车辆所在位置的空驶出租车数比做当前位置的“质量”, 而将两地之间的行驶距离比做万有引力公式中两质量间的距离, 我们就可以构建一个空驶出租车巡游路线选择的引力模型。通过两个位置间“引力”的大小决策最佳的巡游路线。

上述的引力模型可以对空驶出租车在当前时刻选择下一步巡游目的地提供一定的决策依据, 但是未能利用整个出租车服务系统在时间空间上历史演变经验, 具有一定的短视性。因此, 我们在巡游路线的决策中引入增强学习理论, 内化决策的历史经验, 提升短期决策的全局优化特征。

文献[1]指出通过出租车调度优化, 出租车空驶费用可显著下降。文献[2]对出租车的定价范围与车队规模进行了分析。当前关于出租车服务系统的优化研究主要集中在运行模式分析[3] [4] [5]、交通需求的弹性特征分析[6]、市场中存在的竞争特征分析[7]、乘客的随机需求分析[8]、以及电招系统的选择上[9] [10]。近年来, 新的研究方向包括采用新的元胞网络研究出租车网络[11]、考虑需求相应式服务[12], 以及司机与乘客的匹配优化[13]。近年来, 研究者也开始关注机场的出租车调度问题, 以期实现出租车与航班的匹配[14] [15]。文献[16]利用增强学习理论对网格化的出租车实时动态调度进行了建模分析。已有研

究主要关注静态的出租车系统, 一般假设相关的出行需求和车辆巡游速度已知。总而言之, 现有研究对动态的出租车运营系统特征考虑不足, 较难满足今天出租车服务行业的变化需求。

## 2. 基本参变量与系统运行过程描述

### 2.1. 基本参变量

$T$  是调度的实践间隔, 即控制的时间步长。

$k$  是将时间分段后的标记,  $k \in \{0, 1, \dots, K\}$ 。  $t = kT$  是一个时间点,  $KT$  为总的调度时长。

$g \in G$  是一个节点, 即乘客的备选上下车位置;  $G$  是相关的节点集合。

$l \in L$  是一条有向路段, 一般由实际路段和交叉口构成;  $L$  为有向路段集合。

$p \in P$  代表一个乘客,  $P$  是乘客的集合。

$v \in V$  代表一辆出租车,  $V$  是出租车的集合。

$r_g$  是  $T$  内节点  $g$  处乘客的生成率。

$r_{g,h}$  是源于  $g$  处的一个乘客选择节点  $h$  作为目的地的概率。

$\bar{t}_l$  是出租车经过路段  $l$  的平均行程时间。

$\varepsilon_l$  表示通过路段  $l$  时车辆的行程时间误差, 服从均值为 0 而标准方差为  $\delta_l$  的正态分布。

$g_p^O$  是乘客  $p$  的出发点。

$g_p^D$  是乘客  $p$  的达到目的地。

$t_{p,1}$  是乘客  $p$  利用 APP 请求服务的时刻。

$t_{p,2}$  是  $p$  的登车时刻。

$t_{p,3}$  是  $p$  到达目的地时的时刻。

$t_p^W$  表示  $p$  的等车时间。当其出行需求可被出租车满足时,  $t_p^W$  等于  $t_{p,2} - t_{p,1}$ 。

$t_p^R$  是  $p$  的乘行时间。当其需求可被满足时,  $t_p^R$  等于  $t_{p,3} - t_{p,2}$ 。

$d_p$  表示  $p$  的起终点间的乘车距离。

$\eta_v$  为状态变量。当出租车  $v$  有乘客乘坐时,  $\eta_v$  等于 1; 否则, 等于 0。

$g_v$  是  $v$  的标记节点。当有乘客  $p$  乘坐该车辆时,  $g_v$  等于  $g_p^D$ ; 否则, 为空驶出租车  $v$  已选取的最新目的地。

$t_v$  是出租车  $v$  的等待激活时间。当有乘客  $p$  乘坐时,  $t_v$  表示该车距离网格  $g_p^D$  的行程时间; 否则,  $t_v$  为行驶到该车最新目的地的预期行驶时间。

$G_g$  表示与节点  $g$  之间通过一条线段相连的节点集合。

$L_g$  是所有以  $g$  为起点的路段集合。

$P_g(k)$  是  $g$  处在时刻  $kT$  候车的乘客集合,  $n_{g,p}(k)$  是集合  $P_g(k)$  的势;

$V_g^1(k)$  是在时刻  $kT$  在  $g$  处空驶的出租车集合,  $n_{g,v,1}(k)$  是  $V_g^1(k)$  的势。

$V_g^2(k)$  为  $kT$  时刻以  $g$  为目的地的载客出租车集合,  $n_{g,v,2}(k)$  是  $V_g^2(k)$  的势。

$V_g^3(k)$  为  $kT$  时刻以  $g$  为目的地的空驶出租车集合,  $n_{g,v,3}(k)$  是  $V_g^3(k)$  的势。

$c_0$  是乘坐出租车的打车起步价。

$\bar{c}$  表示超出  $d_0$  后, 出租车每公里的单价。

$\hat{t}$  是在起点可接受的最长等车时间。

### 2.2. 系统运行过程

为了简化后续的分析, 我们对出租车服务的路网进行如下的抽象简化。首先, 将现实中常见的出租

车上下客地点, 例如居民小区的出入口、购物中心出租车停靠点、办公大楼出租车停靠点和道路上划定的出租车上下客站点等, 抽象为网络的节点。其次, 将连接上述节点的车辆行驶路径抽象为出租车运行网络的路段。一般而言连接任意两个上下客节点的路径存在多条。我们采用一条抽象的路段代表其中的最短路径, 作为服务网络中节点间的路段。最后, 如果连接路段对应的最短路径需经过其他节点, 则删除该路段。最后的操作可保证网络的节点仅与其附近的乘客上下客节点相连。

下面对出租车服务系统运行过程进行简述。

**步骤 1:** 假设当前时间点为  $kT$ , 当前考虑的节点为  $g \in G$ 。

**步骤 2:** 确定  $P_g(k)$  和  $V_g^1(k)$ 。 $P_g(k)$  由上一控制时刻未得到服务的等待乘客和  $k-1$  时段新生成的乘客在排除超过最长等车时间  $\hat{t}$  的乘客后组成。 $V_g^1(k)$  集合包含了由上一控制时刻在  $g$  处未得到匹配并选择  $g$  继续巡游的空驶出租车、当前时刻  $kT$  在  $g$  处完成载客服的出租车以及在当前时刻  $kT$  抵达  $g$  的空驶出租车。

**步骤 3:** 对空驶出租车  $v \in V_g^1(k)$  与等车乘客  $p \in P_g(k)$  进行司乘匹配。首先对  $P_g(k)$  中的乘客按照等车时间长短进行排序; 其次, 按照排序后的优先次序为乘客指派服务的出租车; 最后, 从  $P_g(k)$  和  $V_g^1(k)$  中删除已经得到匹配的乘客和出租车。此步骤中需记录匹配乘客的等车时间  $t_p^w$  和匹配带来的运营收入, 并更新车辆的特征  $\eta_v$ 、 $g_v$  和  $t_v$ 。

**步骤 4:** 对空驶出租车  $v \in V_g^1(k)$  的巡游路线进行决策。如果没有得到匹配, 空驶出租车可以选择继续在该点巡游或驶向附近的节点。具体的决策方法将在下一节详述。

**步骤 5:** 如果  $G$  中所有节点已经被考虑, 转步骤 6; 否则, 从  $G$  中选取一个未被考虑节点作为新的当前节点, 转步骤 2。

**步骤 6:** 如果  $k < K$ , 令  $k := k+1$ , 并任取节点  $g \in G$ , 更新相关的网络状态变量后转步骤 2; 否则, 终止运行。该步骤中需更新的网络状态变量包括:  $t_p^w$ 、 $t_v$  和与节点相关的集合  $P_g(k)$ 、 $V_g^1(k)$ 、 $V_g^2(k)$  和  $V_g^3(k)$ 。

### 3. 增强学习模型主要组成成分

#### 3.1. 控制行为

在任意给定的控制时刻  $kT$ , 当每一个节点可行的司乘匹配完成后, 可能会出现部分节点还有乘客需要继续等待, 而部分节点出现空驶的出租车。此时, 需要对空驶出租车的巡游路线进行决策。设空载出租车  $v$  所在节点为  $g$ , 那么  $v$  可以选择  $\bar{G}_g \equiv \{g\} \cup G_g$  中的任意节点作为随后的巡游目的地。令  $V(k) := \{v | v \in V, \eta_v(k) = 0\}$ 。对于空驶出租车  $v \in V(k)$ , 其选择的在  $kT$  时刻后的下一巡游目的地节点设为  $h_v(k)$ 。那么由所有  $h_v(k)$ ,  $\forall v \in V(k)$  构成的向量就是对应当前时刻  $kT$  的控制行为, 表示为  $a(k)$  或  $a_k$ 。 $a(k)$  的可行域表示为  $A(k)$ 。

#### 3.2. 控制信息

在增强学习理论中控制信息指的是系统运行中的随机和不确定性因素。出租车服务系统存在两种显著的随机因素。首先, 每个节点乘客的生成是随机的, 具有随机生成率  $r_g$ ,  $\forall g \in G$  和出行目的地分布概率  $r_{g,h}$ ,  $\forall g, h \in G$ 。显然有  $\sum_{h \in G} r_{g,h} = 1$  和  $r_{g,g} = 0, \forall g \in G$  成立。假设  $r_g$  的单位为(人/秒), 仿真试验中我们将逐秒生成(0, 1]间均匀分布的随机数  $RN_1$ 。如果  $RN_1 \leq r_g$ , 则生成一个以该时刻为生成时间的新乘客; 否则, 转下一秒的乘客生成分析。每一个新乘客依据轮盘赌的规则按概率  $r_{g,h}$  选择其目的地。其次, 出租车在路网中两点间的巡游时间具有随机性。受到其他社会车辆和交通控制条件的影响, 出租车的巡游速度具有不确定性。为了模拟上述行程时间的不确定性, 在仿真试验中, 我们假设路网中任意路段  $l$  的行

程时间  $t_i$  由两部分组成, 即  $t_i = \bar{t}_i + \varepsilon_i$ 。为了后续的表达方便, 我们将上述随机因素用抽象向量  $\Delta$  表示。向量  $\Delta$  就是系统所需的控制信息。

### 3.3. 系统状态变量

系统状态变量是给定时刻对系统整体状态加以描述的各种特征因素构成的向量。我们将选择四种系统主要特征构成状态变量。

在给定的控制时间点  $kT$ , 当节点  $g \in G$  处所有的可能司乘匹配完成后, 该点处空驶出租车的数目  $n_{g,v,1}(k)$  和等车乘客数目  $n_{g,p}(k)$  将作为对应点  $g$  的系统状态变量的两个分项。

设有向路段  $l$  的起点为  $g$ , 终点为  $h$ 。起点的  $n_{g,v,1}(k)$  作为该路段起点的引力质量。而在  $kT + \bar{t}_i$  时刻, 完成可能司乘匹配后的  $n_{h,p}(k + \bar{t}_i/T)$  作为该路段终点的引力质量。定义有向路段  $l$  在给定的控制时间点  $kT$  的引力  $f_l(k)$  如下:

$$f_l(k) = \Psi \frac{[n_{g,v,1}(k)]^\alpha [n_{h,p}(k + \bar{t}_i/T)]^\beta}{d_l^2} \tag{1}$$

式(1)中的  $\Psi$ 、 $\alpha$  和  $\beta$  均为给定的正参数。我们选择  $f_l(k)$  作为对应有向路段  $l$  的系统状态变量分项。

一辆空驶的出租车也可能选择当前的节点作停留, 等待可能新出现的服务需求。因此类似于公式(1), 我们定义当前节点  $g$  的引力如下:

$$f_g(k) = \Psi \frac{[n_{g,v,1}(k)]^\alpha [n_{g,p}(k+1)]^\beta}{d_T^2} \tag{2}$$

式(2)中,  $d_T$  表示出租车按照平均巡游速度在单位控制时间段  $T$  的行驶距离。节点的引力  $f_g(k)$  将作为对应点  $g$  的系统状态变量的一个分项。

将所有网络节点上的  $n_{g,v,1}(k)$ 、 $n_{g,p}(k)$  和  $f_g(k)$ ,  $\forall g \in G$  和所有路段上的  $f_l(k)$ ,  $\forall l \in L$  按照一定顺序组成向量  $s(k)$ 。为表述简洁,  $s(k)$  也被简写为  $s_k$ 。  $s(k)$  就是对应控制阶段  $k$  的系统状态变量。所有可行系统状态变量值构成系统状态空间集合  $S$ 。

### 3.4. 控制策略

增强学习理论中, 控制策略指的是由一个给定系统状态变量决定一个控制行为的函数。  $\pi \in \Pi$  表示一个控制策略,  $\Pi$  为所有可行策略集合。控制策略的函数形式可表示为:

$$a = A^\pi(s) \tag{3}$$

增强学习模型的目的之一就是针对给定系统确定可达到某种优化目标的最佳策略。

### 3.5. 行为收益和状态价值

控制行为一旦选定, 将会影响系统当前和随后的运行表现和后续决策。在增强学习理论中, 一般将控制行为对系统当前运行的直接影响定义为控制行为的收益。

假设当前控制时刻为  $kT$ , 节点  $g$  处的一辆空驶出租车  $v \in V_g^1(k)$  选择节点  $h \in \bar{G}_g$  作为其巡游的下一个目的地。为表述方便, 当路段  $l$  以  $g$  为起点而以  $h$  为终点时, 其引力满足  $f_{(g,h)}(k) = f_l(k)$ 。而当  $g$  和  $h$  为同一节点时,  $f_{(g,h)}(k)$  表示节点引力  $f_g(k)$ 。令  $\bar{f}_g(k) := \max_{h \in \bar{G}_g} \{f_{(g,h)}(k)\}$ 。我们定义  $b_{g,h}^v(k) := f_{(g,h)}(k) / \bar{f}_g(k)$  为对应上述目的地选择行为的收益。

对应系统状态变量  $s_k$  和控制行为  $a_k$  的控制行为收益函数定义为:

$$b(s_k, a_k, \Delta_{k+1}) = \left[ \sum_{v \in V(k)} b_{g,h}^v(s_k, a_k, \Delta_{k+1}) \right] / n_{V(k)}. \quad (4)$$

式(4)中的  $\Delta(k+1)$  表示从阶段  $k$  到阶段  $k+1$  系统演变过程中可变为已知的随机控制信息。  $n_{V(k)}$  表示集合  $V(k) := \{v | v \in V, \eta_v(k) = 0\}$  中空驶出租车的数目。

合理的调度控制必须考虑控制行为对出租车服务系统后续表现的影响, 以及这种影响的时间衰减性。因此我们选取如下的优化目标:

$$\max_{\pi \in \Pi} \mathbb{E} \left\{ \sum_{k=0}^{\infty} \left[ \gamma^k b(s_k, A^\pi(s_k), \Delta_{k+1}) \right] \right\}. \quad (5)$$

式(5)中,  $\gamma < 1$  为折扣因子, 算子  $\mathbb{E}$  表示求变量的期望值。式(5)的优化目标为通过对空驶车辆巡游路线的调整, 最大化控制期  $KT$  内阶段性控制行为收益的加权和的期望。

直接求解上述随机优化问题(5)非常困难, 因此我们通常将其转化为等价的 Bellman 方程后再进行求解。设系统处于状态  $s$  时的状态价值为  $V(s)$ 。依据 Bellman 方程, 最优状态价值  $V^*(s)$  应满足:

$$V^*(s_k) = \mathbb{E} \left[ -b(s_k, A^*(s_k), \Delta_{k+1}) + \gamma V^*(s_{k+1} | s_k, A^*(s_k), \Delta_{k+1}) \right]. \quad (6)$$

由式(4)和(5)可知, 状态价值  $V(s)$  的含义为: 考虑时间衰减因素时, 以  $s$  作为起始系统状态时系统在随后的运行过程中可以得到的总的期望收益的相反数。

下节将设计对应的 Q 学习算法求解方程(6)。  $V(s)$  与 Q 函数  $Q(s, a)$  具有如下关系:

$$V(s) = \max_a Q(s, a). \quad (7)$$

与  $V(s)$  相比,  $Q(s, a)$  的变量维度增加了一位; 但是实际应用时可通过比较不同控制行为的  $Q(s, a)$  值更方便地确定最佳控制行为。由于  $Q(s, a)$  的具体形式未知, 因此需要利用某种带有待定参数  $\theta$  的近似函数  $Q(s, a, \theta)$  来替代  $Q(s, a)$  实现问题的求解。

#### 4. 求解算法

下面将介绍求解问题(6)的 Q 学习算法。该方法中的 Q 函数为人工神经网络(ANN)近似函数, 具有待定参数  $\theta$ 。算法中  $k$  代表当前阶段的序列号。  $m$  表示对系统进行仿真模拟的序号。  $K$  和  $M$  分别是总的阶段数目和模拟仿真总次数。针对出租车调度优化的 Q 算法的求解步骤如下:

**步骤 1:** 初始化。

步骤 1.1. 对所有的系统状态  $s(k)$ , 给出价值函数  $\bar{Q}_k^0(s_k, a_k, \theta)$  的近似值以及控制行为  $a(k) \in A(k)$ ,  $k = \{0, 1, \dots, K-1\}$ 。其中  $\theta$  的初始值从区间  $[0, 1]$  中随机生成。

步骤 1.2. 令  $m = 1$ 。

步骤 1.3. 初始化  $s^1(k)$ 。

**步骤 2:** 选择一个随机信息的样本路径  $\omega^m$ 。

**步骤 3:** 将下面操作步骤依  $k = 0, 1, \dots, K-1$  加以循环迭代:

步骤 3.1. 利用  $\varepsilon$  贪婪规则确定控制行为。以概率  $\varepsilon$ , 从控制行为集  $A(k)$  中随机选择行为  $a_k^m$ 。而以概率  $1 - \varepsilon$ , 利用公式  $a_k^m \in \arg \max_{a_k \in A(k)} \bar{Q}_k^{m-1}(s_k^m, a_k, \theta)$  选择行为  $a_k^m$ 。

步骤 3.2. 对信息  $\Delta_{k+1}^m = \Delta(s_k^m, a_k^m, \Delta_{k+1}^m)$  进行取样, 并计算下一阶段的状态变量  $s_{k+1}^m = S^\Gamma(s_k^m, a_k^m, \Delta_{k+1}^m)$ 。  $S^\Gamma(\cdot)$  为系统状态转移函数, 表示系统演进到下一控制状态。

步骤 3.3. 计算  $\hat{Q}_k^m = -b(s_k^m, a_k^m, \Delta_{k+1}^m) + \gamma \max_{a_{k+1} \in A(k+1)} \bar{Q}_{k+1}^{m-1}(s_{k+1}^m, a_{k+1}, \theta)$ 。

步骤 3.4. 更新  $\bar{Q}_k^m$  和  $\theta$ 。

首先, 依据下式对  $\theta$  进行更新:

$$\theta := \theta + \lambda \nabla \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta) (\hat{q}_k^m - \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta)).$$

其次, 基于更新后的参数  $\theta$ , 对 Q 因子进行如下更新:

$$\bar{Q}_k^m(s_k^m, a_k^m) = (1 - \alpha_{m-1}) \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta) + \alpha_{m-1} \hat{q}_k^m.$$

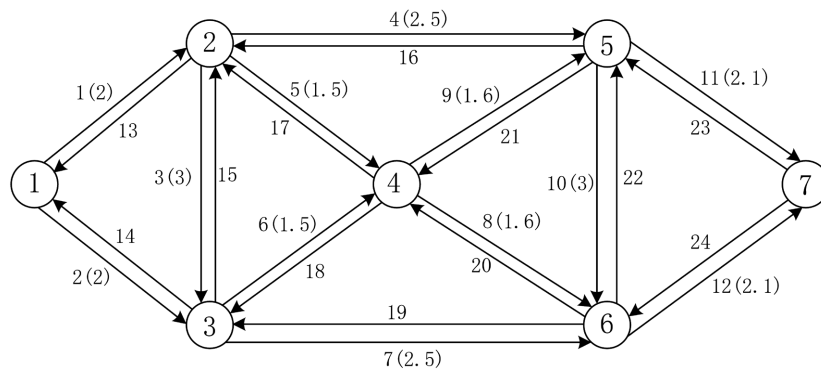
**步骤 4:** 令  $m := m + 1$ 。如果  $m \leq M$ , 转步骤 2。

**步骤 5:** 给出最终的 Q 因子  $(\bar{Q}_k^M)_{k=1}^{K-1}$  和  $\theta$ 。

$\nabla \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta) (\hat{q}_k^m - \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta))$  可看作函数  $z(\theta) = (1/2) (\hat{q}_k^m - \bar{Q}_k^{m-1}(s_k^m, a_k^m, \theta))^2$  的负梯度方向。这里  $\theta$  作为自变量, 而  $\hat{q}_k^m$  为给定值。那么上述算法可以看作是在求解  $z(\theta)$  的最小值。依据经典非线性规划理论, 如果当前  $\theta$  对应的函数值非局部或全局最小值, 且步长足够小, 目标函数  $z(\theta)$  的值将沿当前变量  $\theta$  处的负梯度方向下降。因此算法中的参数  $\lambda$  相当于沿可行下降方向的搜索步长。

### 5. 算例分析

考虑图 1 所示的出租车服务运行网络。路段的标号已在图中标出, 路段标号后括号内的数字表示该路段的长度(km)。假设连接两个节点的两条相反方向的路段长度相等。假设出租车在所有路段上的平均巡游速度均为 10 m/s (36 km/hr)。由各路段长度除以该平均速度即可得到每条路段的平均行程时间。例如路段 4 的平均行程时间为  $2500 \text{ (m)}/10 \text{ (m/s)} = 250 \text{ (s)}$ 。通过路段  $l$  的行程时间误差  $\varepsilon_l$  服从正态分布  $N(0, 5d_l)$ , 其中  $d_l$  为路段  $l$  的公里数。



**Figure 1.** The operation network of taxi service  
**图 1.** 出租车服务运行网络

**Table 1.** The generation rate of passengers  $r_g$  and the gathering coefficient  $\varpi_g$

**表 1.** 节点的乘客生成率  $r_g$  和乘客集聚系数  $\varpi_g$

节点	1	2	3	4	5	6	7
$r_g$	0.2	0.3	1.3	0.2	0.5	0.2	0.5
$\varpi_g$	0.5	1.2	0.2	0.2	0.5	2.1	1.2

表 1 给出了节点的乘客生成率  $r_g$  和乘客集聚系数  $\varpi_g$ 。此表中  $r_g$  的单位为(乘客数/分钟)。这里乘客集聚系数可以理解为该节点作为出租车乘客目的地的相对吸引力, 无量纲。利用乘客集聚系数  $\varpi_g$  可以通过下面的式(8)和式(9)得到乘客的目的地选择概率  $r_{g,h}$ :

$$r_{g,h} = (\varpi_h / \bar{t}_{g,h}) / \left[ \sum_{j \in G/\{g\}} (\varpi_j / \bar{t}_{g,j}) \right], \quad \forall h \in G/\{g\}, \quad (8)$$

$$r_{g,g} = 0. \quad (9)$$

网络中共有 20 辆出租车。在调度开始时, 出租车均为空驶状态, 且随机分布在 7 个节点处。设总的时间分段数  $K$  为 100, 总的仿真次数  $M$  为 300, 参数  $\varepsilon$  和折扣因子  $\gamma$  均设为 0.5, 算法步长  $\lambda$  设为 0.05。引力公式(1)和(2)中的参数  $\alpha$  和  $\beta$  取值均为 1, 参数  $\Psi$  取值为 100。其他的系统参数为  $T = 100 \text{ s}$ 、 $c_0 = 14 \text{ 元}$ 、 $\bar{c} = 2.5 \text{ 元}$ 、 $d_0 = 3 \text{ km}$  和  $\hat{t} = 400 \text{ s}$ 。

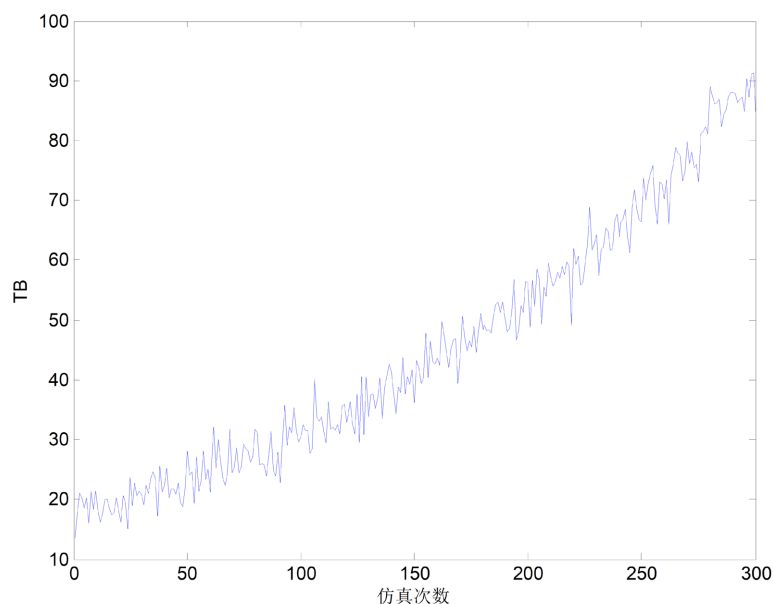
表 2 中给出了控制与无控制情景下常见的出租车服务水平度量指标, 包括出租车总的空驶里程(Total travel Distance of Vacant taxis-TDV)、总的司乘匹配数(Number of passenger and taxi Matching Pairs-NMP)、总的出租车运营收入(Total Income of Operation of taxis-TIO)、未得到服务而损失的乘客数(Un-Served Passenger-USP)以及一次仿真过程中所有控制行为的总控制收益(Total Benefit-TB)。表中数据表明服务水平指标在调度优化控制条件下均有明显的改善。因为系统本身的随机特征, 上述数据和随后给出的数据均为系统经过一次仿真或多次仿真得到的代表性结果。

**Table 2.** The comparison of measure indexes of service level

**表 2.** 出租车服务水平度量指标比较

情景	TB	TDV (km)	NMP	USP	TIO (元)
无控制	24.58	801.5	345	186	5016.0
调度控制	91.02	588.3	446	67	6392.2

在图 2、图 3 和图 4 中, 我们分别给出了总的控制收益(TB)、总的空驶里程(TDV)以及总的出租车运营收入(TIO)随仿真次数增加的变化情况。尽管上述三个指标由于系统本身随机特性而具有强烈的波动特征, 但是三个指标的总体变化趋势均呈现提升特征。总控制收益(TB)和总运营收入(TIO)表现出上升趋势, 而总空驶里程(TDV)表现出下降趋势。



**Figure 2.** The change trend of TB

**图 2.** 仿真中总的控制收益(TB)变化



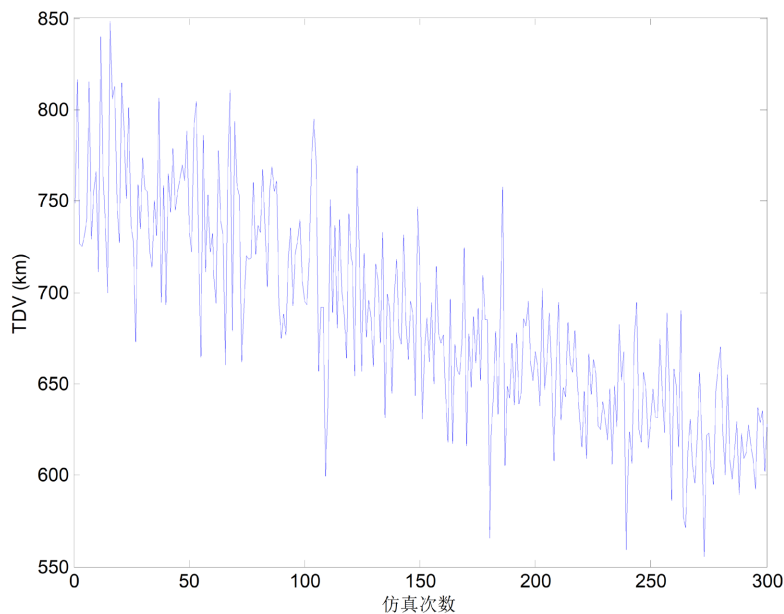


Figure 3. The change trend of TDV  
图 3. 出租车总的空驶里程(TDV)变化

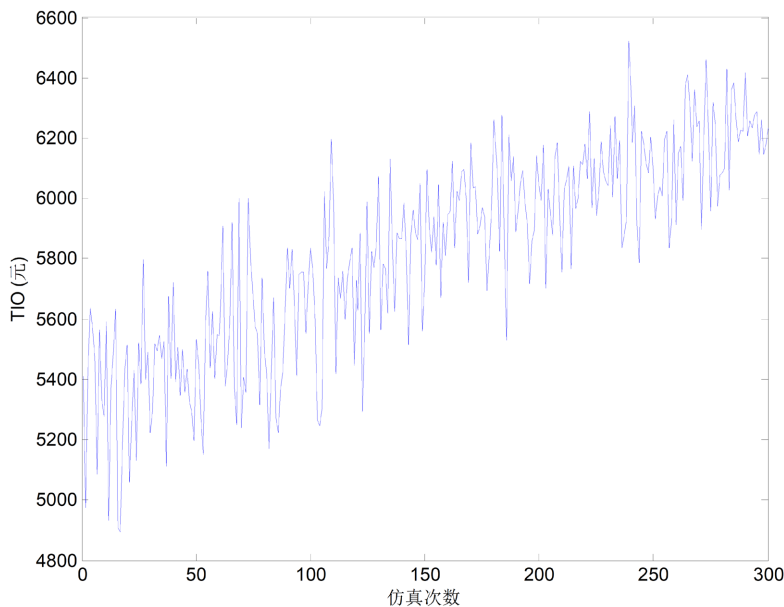


Figure 4. The change trend of TIO  
图 4. 总的出租车运营收入(TIO)变化

图 5 给出了总的司乘匹配数(NMP)和未得到服务损失乘客数(USP)随仿真次数增加的变化趋势。总的司乘匹配数(NMP)呈现上升趋势, 而未得到服务损失乘客数(USP)呈现下降趋势。值得注意的是给定仿真下两者和的均值为 528.6567, 上下变化的区间为[470, 590]。由表 2 中的乘客生成率  $r_g$  (乘客数/分钟), 可知  $\sum_g (r_g / 60)$  为系统一秒钟内生成的乘客数均值。而总的控制时长为  $KT$  (秒), 因此在控制时长内生成的期望乘客数为  $KT \sum_g (r_g / 60) \approx 533$ 。该期望乘客数与 300 次仿真中生成的乘客均值一致。

算例求解的计算机程序用 Java 1.8.0 编写, 在 NetBeans IDE 8.0.2 开发环境下实现, 所用计算机处理

器为 Intel® Core i3-3120M CPU。调度控制方式下完成 300 次系统仿真的计算时间约为 8 s, 而完成一次仿真的计算时间变化范围为[0.014 s, 0.032 s]。

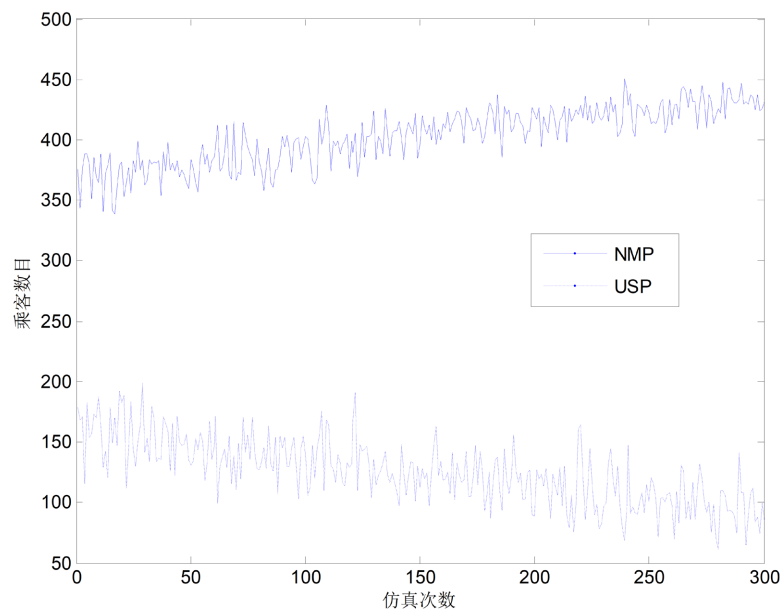


Figure 5. The change trends of NMP and USP

图 5. 总的司乘匹配数(NMP)和未得到服务损失乘客数(USP)的变化

## 6. 结论

通过类比万有引力定理, 将出租车的揽客行为加以建模, 并从时空演化的角度引入增强学习理论, 从而实现城市规模的区块化出租车动态调度与空驶路线调整。研究表明: 为增加揽客成功的机率, 不仅需要考虑到预期的目的地等车乘客数量, 也需考虑到到达该目的地的行程距离带来的不确定性风险; 出发点空驶出租车数辆, 也可对司乘的最终匹配结果产生影响; 优化后, 出租车总的空驶里程和总的运营收入都有显著增加, 总的司乘匹配数也有所增加, 而未得到服务损失乘客数有所下降。

## 基金项目

国家自然科学基金项目(71601118, 71801153, 71871144); 上海市自然科学基金项目(18ZR1426200)。

## 参考文献

- [1] Zhan, X., Qian, X. and Ukkusuri, S.V. (2016) A Graph-Based Approach to Measuring the Efficiency of an Urban Taxi Service System. *IEEE Transactions on Intelligent Transportation Systems*, **17**, 2479-2489. <https://doi.org/10.1109/TITS.2016.2521862>
- [2] Beesley, M.E. (1973) Regulation of Taxis. *The Economic Journal*, **83**, 150-172. <https://doi.org/10.2307/2231106>
- [3] Yang, H. and Wong, S. (1998) A Network Model of Urban Taxi Services. *Transportation Research Part B*, **32**, 235-246. [https://doi.org/10.1016/S0191-2615\(97\)00042-8](https://doi.org/10.1016/S0191-2615(97)00042-8)
- [4] Jung, J., Chow, J.Y., Jayakrishnan, R. and Park, J.Y. (2014) Stochastic Dynamic Itinerary Interception Refueling Location Problem with Queue Delay for Electric Taxi Charging Stations. *Transportation Research Part C*, **40**, 123-142. <https://doi.org/10.1016/j.trc.2014.01.008>
- [5] Sayarshad, H.R. and Chow, J.Y. (2016) Survey and Empirical Evaluation of Nonhomogeneous Arrival Process Models with Taxi Data. *Journal of Advanced Transportation*, **50**, 1275-1294. <https://doi.org/10.1002/atr.1401>
- [6] Wong, K., Wong, S. and Yang, H. (2001) Modeling Urban Taxi Services in Congested Road Networks with Elastic

- Demand. *Transportation Research Part B*, **35**, 819-842. [https://doi.org/10.1016/S0191-2615\(00\)00021-7](https://doi.org/10.1016/S0191-2615(00)00021-7)
- [7] Yang, H., Wong, S.C. and Wong, K. (2002) Demand-Supply Equilibrium of Taxi Services in a Network under Competition and Regulation. *Transportation Research Part B*, **36**, 799-819. [https://doi.org/10.1016/S0191-2615\(01\)00031-5](https://doi.org/10.1016/S0191-2615(01)00031-5)
- [8] Zhang, W. and Ukkusuri, S.V. (2016) Optimal Fleet Size and Fare Setting in Emerging Taxi Markets with Stochastic Demand. *Computer-Aided Civil and Infrastructure Engineering*, **31**, 647-660. <https://doi.org/10.1111/mice.12203>
- [9] Wong, K., Wong, S., Yang, H. and Wu, J. (2008) Modeling Urban Taxi Services with Multiple User Classes and Vehicle Modes. *Transportation Research Part B*, **42**, 985-1007. <https://doi.org/10.1016/j.trb.2008.03.004>
- [10] He, F. and Shen, Z.-J.M. (2015) Modeling Taxi Services with Smart Phone-Based e-Hailing Applications. *Transportation Research Part C*, **58**, 93-106. <https://doi.org/10.1016/j.trc.2015.06.023>
- [11] Wong, R., Szeto, W. and Wong, S. (2014) A Cell-Based Logit-Opportunity Taxi Customer-Search Model. *Transportation Research Part C*, **48**, 84-96. <https://doi.org/10.1016/j.trc.2014.08.010>
- [12] Amirgholy, M. and Gonzales, E.J. (2016) Demand Responsive Transit Systems with Time-Dependent Demand: User Equilibrium, System Optimum, and Management Strategy. *Transportation Research Part B*, **92**, 234-252. <https://doi.org/10.1016/j.trb.2015.11.006>
- [13] Wang, X., He, F., Yang, H. and Gao, H.O. (2016) Pricing Strategies for a Taxi-Hailing Platform. *Transportation Research Part E*, **93**, 212-231. <https://doi.org/10.1016/j.tre.2016.05.011>
- [14] Xu, J., Rahmatizadeh, R., Bölöni, L., et al. (2021) A Taxi Dispatch System Based on Prediction of Demand and Destination. *Journal of Parallel and Distributed Computing*, **157**, 269-279. <https://doi.org/10.1016/j.jpdc.2021.07.002>
- [15] Wang, Y. (2020) Research on Airport Taxi Dispatching Based on Probability Model. *Journal of Artificial Intelligence Practice*, **3**, 59-62.
- [16] 何胜学. 基于增强学习的网格化出租车调度方法[J]. 计算机应用研究, 2019, 36(3): 762-766.