

doi: 10.3969/j.issn.1002-0268.2021.04.016

基于双边理论的顺风车稳定匹配优化

马瑞民¹, 姚立飞²

(1. 广州大学 管理学院, 广东 广州 510006; 2. 广东财经大学 地理与旅游学院, 广东 广州 510320)

摘要: 智能手机、移动平台和在线支付系统的广泛使用为顺风车出行提供了新的机会, 顺风车是一种持续的交通方式, 可以充分利用交通系统中现有私家车辆的空闲承载能力。该文引入稳定性的概念来建立一个顺风车稳定匹配模型, 并提出一种启发式算法来找到稳定或近似稳定的匹配方案, 并进行数值模拟来证明所提出算法的计算效率, 探索该方法在一定规模下的顺风车匹配的实际适用性。结果表明: (1) 随着参与者数量的增加, 匹配率 Suc 、总距离节省比率 Sav 和参与者个人出行距离节省比率 $Sipr$ 都不断上升, 且有两个交通聚集中心时, Suc , Sav 和 $Sipr$ 都比没有交通中心时高, 而司机绕路比率 Di 不断下降; (2) 随着参与者时间机会成本系数的增加, 匹配率 Suc 、总距离节省比率 Sav 和司机绕路比率 Di 不断下降; (3) 随着参与者弹性时间的增加, 匹配率 Suc 、总距离节省比率 Sav 和参与者个人出行距离节省比率 $Sipr$ 都上升, 司机绕路比率 Di 下降; (4) 通过比较系统最优化模型, 考虑参与者成本的系统优化模型和稳定匹配优化模型求解时间, 发现所提出的偏好列表精简算法能够有效地降低求解稳定匹配解的时间。

关键词: 智能交通; 顺风车出行; 稳定匹配; 启发式算法; 稳定性

中图分类号: U491

文献标识码: A

文章编号: 1002-0268 (2021) 04-0131-11

Optimization of Ride-sharing Stable Matching Based on Two-side Theory

MA Rui-min¹, YAO Li-fei²

(1. School of Management, Guangzhou University, Guangzhou Guangdong 510006, China;

2. School of Geography and Tourism, Guangdong University of Finance and Economics, Guangzhou Guangdong 510320, China)

Abstract: The wide-spread use of smart phones, mobile platforms and online payment systems provides new opportunities to enable ridesharing, which is a continuous transport mode that can make full use of the free carrying capacity of existing private vehicles in the transport system. A notion of stability is introduced to build a stable ridesharing matching model. A heuristic algorithm is proposed to find a stable or approximately stable matching scheme, and numerical simulation is carried out to prove the computational efficiency of the proposed algorithm, and to explore the practical applicability of the method in a certain scale of ridesharing matching. The result shows that (1) With the increasing of number of participants, the matching rate Suc , the total distance saving rate Sav and the individual travel distance saving rate $Sipr$ increase. When there are 2 traffic centers, Suc , Sav and $Sipr$ are respectively higher than those when there are no traffic center, while the driver detour ratio Di decreases. (2) With the increase of time opportunity cost coefficient of participants, Suc , Sav and Di decrease. (3) With the increase of participants' elastic time, Suc , Sav and $Sipr$ increase, while Di decreases. (4) By comparing the solution time of the system optimization model, the system optimization model considering the cost of participants and the stable matching optimization model, it is found that the proposed preference list reduction algorithm can effectively reduce the solution time of stable matching.

收稿日期: 2020-06-10

基金项目: 教育部人文社会科学项目 (19YJC630119)

作者简介: 马瑞民 (1989-), 男, 湖南隆回人, 博士. (maruimin@gzhu.edu.cn)

Key words: ITS; ridesharing; stable match; heuristic algorithm; stability

0 引言

私人小汽车保有量持续增加,座位占用率却一直很低。截至2019年底,我国私人汽车拥有量达2.32亿辆,其中载客汽车1.89亿辆,但从平均保有量来讲相比美国等发达国家仍然较低,这意味着我国私人汽车保有量很有可能将会持续增加。但是,目前私人小汽车出行的座位占用率仍然很低。据调查北京市在上下班高峰期80%的私家车上都只有一个人^[1],整个交通网络中的私人汽车座位的使用率极低。由此可见,如果能够充分挖掘这部分交通网络中尚未使用交通运输能力,将对整个交通系统产生巨大的作用,而这也正是顺风车的兴起的契机。

移动互联平台、智能手机及GPS等技术的普及为顺风车提供了新的机遇,目前已经有多个平台提供顺风车出行服务,如美国的UberX和Lyft,欧洲的BlaBlaCar以及国内的滴滴出行和嘀嗒出行等等。顺风车出行是一种灵活的交通方式,平台中司机和乘客给出出发时间及起止点,有平台自动对其进行匹配并具有固定的费用计算规则,一旦匹配成功平台将对司乘双方告知信息。顺风车出行与传统的出租车出行是不同的,顺风车出行不以盈利为目的,旨在司乘双方的出行费用节省,且顺风车出行中司乘双方都是平台的顾客,双方都可根据其偏好自主选择一起出行。司机与乘客之间的匹配是顺风车出行中的一个关键问题。大多数研究将问题表述为具有不同系统优化目标的集中分配问题,如最小化总运行距离^[2-4]、最小化总运行时间^[5]、最大化系统匹配数量^[6-7]等,且系统最优化匹配求解方法也是多种多样,有精确算法^[8-9]、遗传算法^[10]、贪婪算法^[11]以及蜂群算法^[12]等。但系统最优并不一定是参与者个人的最佳选择,在顺风车出行平台中参与者是独立的、利己的个体,如果他们相信存在更好的匹配,他们将拒绝平台给出的匹配方案,接着去寻找更好的匹配对象或直接转向其他的出行平台,即平台给出的匹配方案并不稳定。

稳定匹配的一个基本问题是众所周知的稳定婚姻问题(Stable Marriage Problem, SMP),它是由Gale和Shapley在1962年的开创性论文中首次提出的^[13]。针对SMP提出了一种延迟接受算法(DA),该算法可以为任何稳定的婚姻问题找到至少一个稳定的匹配解。然后,越来越多的SMP变体被提

出^[14]。在本研究中,往往存在参与者中的某一方因时间窗约束不接受方案或因成本分配不接受方案,而这些参与者是不会出现在偏好列表中的,即在顺风车匹配问题中偏好列表是不完全的。此外,参与者在不同的匹配中可能获得相同的利益,并不严格地偏好其中一种匹配。综上,本研究将从参与者个体角度出发,引入稳定性概念,构建司乘双方稳定匹配模型,并提出相对应的启发式算法提高模型求解速度。

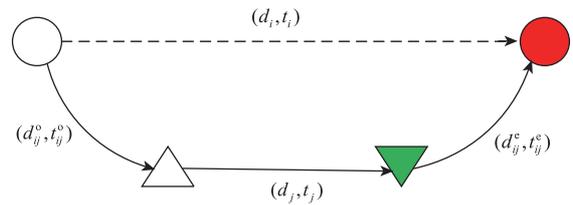


图1 司机*i*与乘客*j*之间的拼车出行

Fig. 1 Carpooling between driver *i* and passengers *j*

1 问题描述

一些学者已经构建了相关的顺风车匹配系统最优化模型^[2]:

$$f = \max \sum_{i=1}^m \sum_{j=1}^n x_{ij} \cdot s_{ij}, \quad (1)$$

$$\text{s. t. } \max(ed_i + t_{ij}^o, ed_j) \leq \min(la_i - t_{ij}^e - t_j, la_j - t_j), \quad (2)$$

$$\sum_{i=1}^n x_{ij} \leq 1, \quad \forall j, \quad (3)$$

$$\sum_{j=1}^m x_{ij} \leq 1, \quad \forall i, \quad (4)$$

$$x_{ij} \in \{0, 1\}, \quad \forall i, j, \quad (5)$$

式(1)为目标函数,其中 $s_{ij} = (d_i + d_j) - (d_{ij}^o + d_j + d_{ij}^e)$ 表示司机*i*与乘客*j*拼车后所贡献的距离节省, d_i, d_j 分别表示司机*i*和乘客*j*起点到终点的原始距离; d_{ij}^o 为司机*i*起点到乘客*j*起点的距离; d_{ij}^e 为乘客*j*终点到司机*i*终点的距离。式(2)为时间窗约束,每一对可行的匹配之间肯定存在以下的时间约束:(1)司机的最大可用时间必须不能小于完成合乘路线所需要的总时间;(2)司机行驶到乘客起点的时间窗与乘客出发的时间窗是有交叉的,即至少存在一个时间点能够满足司机顺利接到乘客,其中 t_i 和 t_j 分别为司机*i*和乘客*j*起点到终点的原始时间, t_{ij}^o 为司机*i*起点到乘客*j*起点的的时间, t_{ij}^e 为乘客*j*终点到司

机 i 终点的时间, ed_i, la_i 和 ed_j, la_j 分别表示司机 i 和乘客 j 最早出发时间和最晚到达时间。式 (3) ~ (5) 为 0-1 变量约束, 该模型研究的是单司机单乘客匹配模型, 即一个司机在同一时间只能搭载一个乘客, 一个乘客在同一时间也只能乘坐一个司机的车。

上述模型旨在实现系统最优化的目标, 即系统中所有车辆的总行驶距离最小化或系统中所有车辆总行驶时间最小化, 但并未考虑匹配双方的成本分配, 在系统最优模型的方案中有可能出现司机的匹配出行的成本还高于单独出行的成本的情况。从参与者个体角度来看, 系统最优化模型给出的方案对参与者来说并不是最佳选择的选择。

2 顺风车出行稳定匹配模型

首先建立稳定匹配问题的线性规划模型, 并找出一个最优的稳定匹配解, 记为 f_{sta} 。对于平台而言, 获利来源于匹配成功之后的距离节省 s , 假设平台从每一对成功匹配都得到一个固定比例的利润, 令平台获利比例为 η 。

2.1 模型假设

A1: 一个司机只能载一个乘客, 一个乘客也只能搭乘一个司机的车。

A2: 乘客与司机的出行成本完全取决于车辆的行驶距离。

A3: 如果参与者没有找到合适的匹配对象, 将独自开车出行。

A4: 参与者对成本理性, 所有可行的匹配都是对参与者有成本节省的。

A5: 参与者知道所有潜在匹配对象的信息。

2.2 参与者偏好序

(1) 参与者效用函数构建

在本研究中共享出行参与者的成本分为两个部分: 距离成本和时间机会成本。

假设 (i, j) 是最优稳定解中的一个匹配对, 那么 (i, j) 的距离节省为 s_{ij} , 记平台获利为 φ_{ij} , 那么平台因为司机 i 和乘客 j 匹配的获利为 $\varphi_{ij} = \alpha \cdot \eta \cdot s_{ij}$ 。令司机 i 的时间机会成本为 $wt_i = \omega \cdot (t_{ij}^o + t_j + t_{ij}^e - t_i)$, 乘客 j 的时间机会成本为 $wt_j = \omega \cdot \max(0, ed_i + t_j - ed_j)$, 其中 ω 表示单位时间的工资水平, 令 $u_i(i, j), u_j(i, j)$ 分别代表匹配对 (i, j) 中司机 i 与乘客 j 的成本效用值, 那么效用值计算式可以表示为:

$$u_i(i, j) = \alpha \cdot s_{ij} + Pay(i, j) - wt_i, \quad (6)$$

$$u_j(i, j) = \alpha \cdot d_j - Pay(i, j) - wt_j, \quad (7)$$

式中 α 表示单位距离的成本; $Pay(i, j)$ 为乘客 j 向司机 i 支付的车费。而 $Pay(i, j)$ 的确定必须考虑参与者之间的具体成本分配方式, 不同的分配方式 $Pay(i, j)$ 是不一样的。

目前关于单司机单乘客匹配模型中成本分配机制主要分为 4 种: (1) Geisberger 等 (2009) 提出平均分配司机与乘客合乘段的成本^[15]; (2) Kleiner 等 (2011) 提出平均分配总的行驶距离^[16]; (3) Wang 等 (2013) 提出平均分配可行匹配的成本节省^[17]; (4) Agtaz 等 (2011) 按原始距离比例进行参与者之间的成本分配^[2]。采用原始距离比例进行参与者之间的成本分配, 则有:

$$Pay(i, j)_{gen} = \alpha \cdot [d_j - \bar{\omega}(1 - \eta)s_{ij}], \quad (8)$$

式中 $\bar{\omega}$ 为成本节省的分配比率, $\bar{\omega} = d_j / (d_i + d_j)$ 。

(2) 参与者效用值排序

由于参与者是理性的, 即拼车出行成本必须小于自己驾驶出行的成本, 令自己出行的效用值 $u(p, p) = 0$, 即若参与者没有因为参与到共享出行平台而有所改变或获得利益, 参与者将自己出行。

所有参与者最终有两种匹配结果, 第 1 种是存在可行的匹配对象并且匹配成功, 此时效用值 $u_i(i, j) > 0$ 且 $u_j(i, j) > 0, \forall k \in I, l \in J$; 第 2 种是有可接受的匹配对象但是没有匹配成功或者没有接受的匹配对象, 此时效用值不大于 0, 即 $-(u_i(i, j) > 0 \& u_j(i, j) > 0), \forall k \in I, l \in J$, 这一类匹配对象将会直接从该参与者的偏好序中删除。那么此时视为参与者与自己进行匹配。下面将参与者的偏好序定义为:

$$\pi(p) = ([1_l], [2_k], \dots, [n_p], p), p \in P. \quad (9)$$

$\pi(p)$ 是根据参与者 p 的效用值进行排序得到的一个偏好序列, 下面以具体例子具体说明偏好序中各个符号的意义。例如用 $\pi(j) = ([1_l], [2_k], \dots, [n_i], j)$ 为乘客 j 的偏好序, 那么 $[1_l]$ 为司机 $l \in I$ 处在乘客 j 偏好序的首位, 即对乘客 j 来说, 司机 l 就是其最佳选择; 其中元素 $[n_i]$ 表示潜在的可接受的匹配对象处在列表 $\pi(j)$ 的第 n 个位置, n 也表示可行的匹配对象的数量。在偏好序的最后一个位置是参与者自己, 表示如果参与者没有找到可行的匹配对象, 那么他将独自开车出行或继续留在平台上等待匹配, 直到他发布的出行信息过期 (超出预计的出发时间)。令 $j >_i j'$ 表示司机 i 严格偏好乘客 j 于 j' , 同理, $i >_j i'$ 表示司机 j 严格偏好乘客 i 于 i' 。如果司机 i 与乘客 j 之间的偏好关系满足 $i >_j j$, 那么对乘客 j 来说司机 i 是一个潜在的匹配对象, 同理,

$j >_i i$ 表示乘客 j 是司机 i 的潜在接匹配对象。如果司机 i 与乘客 j 之间的偏好关系满足 $i >_j j$ 且 $j >_i i$, 那么 (i, j) 是一个可行的匹配对。

由上述分析我们可以对偏好序中的参与者分为两类, 第1类是存在可行的匹配对象并且匹配成功, 这一类的司机与乘客分别记为 $I(j) \in I$ 和 $J(i) \in J$; 第2类是有可接受的匹配对象但是没有匹配成功或者没有接受的匹配对象, 这一类分别的匹配对象就是参与者自己, 记为 $\{i\}$ 和 $\{j\}$ 。令 $M = I(j) \cup \{j\}$ 表示乘客 $j, j \in J$ 的策略空间, 令 $W = J(i) \cup \{i\}$ 表示司机 $i, i \in I$ 的策略空间。令 ψ 表示可行匹配对 (i, j) 的集合, 其中 $i \in M, j \in W$ 。

从上述分析中可以知道, 对于任何一个匹配对 $(i, j) \in \psi$ (其中 $i \in M, j \in W$) 都是可行的, 也就是说 (i, j) 满足了对应的时间窗口约束, 同时考虑了司乘双方的成本分配问题, (i, j) 也能满足匹配后的出行成本小于或等于匹配前的出行成本的成本约束。于是, 可以将系统优化模型式 (1)~(5) 转化为:

$$f_{so} = \max \sum_{(i,j) \in \psi} x_{ij} \cdot s_{ij}$$

$$\text{s. t. } \begin{cases} \sum_{i \in M} x_{ij} \leq 1, \forall j \\ \sum_{j \in W} x_{ij} \leq 1, \forall i \\ x_{ij} \in \{0, 1\}, \forall i, j \end{cases}, \quad (10)$$

式中 f_{so} 是考虑参与者之间成本分配后的系统最优值, 模型 (10) 能够保证参与者之间成本分配后个人成本节省是非负的, 即保证了出行成本会有减少, 但是并不能解决前文所提到的稳定匹配。

2.3 稳定优化模型

对于参与者来说, 系统可能会存在一些无差异偏好序列, 即对着某个参与者来说, 存在多个匹配对象的偏好一样。如此一来, 这个参与者将有一个不严格偏好序, 为了解决这个问题, Vate 和 Roth 等提出的线性规划方法对该问题进行求解。稳定匹配模型如下^[18-19]:

$$f_{sta} = \max \sum_{i \in M} \sum_{j \in W} x_{ij} (d_i - d_{ij}^o - d_{ij}^e)$$

$$\begin{cases} \sum_{i \in M} x_{ij} = 1, \forall j \in J \\ \sum_{j \in W} x_{ij} = 1, \forall i \in I \\ \sum_{i \geq j} x_{ij} + \sum_{j \geq i} x_{ij} + x_{ij} \geq 1, \forall i \in M, j \in W \\ x_{ij} \geq 0, \forall i \in M, j \in W \end{cases}. \quad (11)$$

对于小规模匹配问题求解稳定最优解时, 可以采用 CPLEX 进行求解, 但是存在以下两个问题: (1) 本研究要处理的是顺风车共享出行参与者之间的匹配, 涉及的参与者数量会比较大, 而且要求匹配迅速及时。(2) Vate 和 Roth 的优化模型中, 仍然以系统最优化为目标, 并没有考虑参与者个人的满意程度。因此, 要将 Vate 和 Roth 等提出的线性规划方法应用到现实, 就只能考虑将该线性规划模型的可行域空间尽可能地缩小, 这正是本研究研究的重点问题。

3 模型求解

3.1 可行域空间缩减方法

当参与者的数量较小时, 可以对模型 (11) 直接进行最优解求解, 但是本研究需要处理的是共享出行平台较大批量的司机与乘客的出行匹配, 需要短时间进行大批量的匹配, 如果不能对模型的可行域进一步地缩减, 很难满足现实中的需求。

3.1.1 偏好列表相关理论

定义 1^[19]: 一对一的司机乘客匹配可以定义为从参与者全集到参与者全集的一种单射关系, $\mu: M \cup W \rightarrow M \cup W$, 对任意的 $i \in M$ 和 $j \in W$, 满足以下关系:

- (1) $\mu(i) = j$ 当且仅当 $\mu(j) = i$, 此时 i 和 j 是匹配对, 记为 (i, j) ;
- (2) 如果 $\mu(i) \notin M$, 则 $\mu(i) = i$, 此时 i 是没有匹配上的, 记为 (i, i) ;
- (3) 如果 $\mu(j) \notin W$, 则 $\mu(j) = j$, 此时 j 是没有匹配上的, 记为 (j, j) 。

其中, 根据 μ 确定的所有匹配对的集合称为匹配方案 μ 。

定义 2: 一对一双边匹配 $\mu: M \cup W \rightarrow M \cup W$, 对任意的 $i \in I$ 和 $j \in J$, 如果 i 和 j 满足 $i >_{\mu} j$ 且 $j >_{\mu} i$, 则称 (i, j) 是匹配方案 μ 的一个阻碍对。

定理 1 假设 γ 和 γ' 表示偏好结构, 且满足 $\gamma' \pm_r \gamma$ 。令 μ 和 μ' 表示相对应的稳定匹配, $M_{\mu} (M_{\mu}')$ 表示偏好 μ 于 μ' (偏好 μ' 于 μ) 的司机的集合, 同样的道理, 令 $W_{\mu} (W_{\mu}')$ 表示偏好 μ 于 μ' (偏好 μ' 于 μ) 的乘客的集合。那么, μ' 和 μ 与 M_{μ}' 和 W_{μ} 之间是一种双射关系。

证明: 假设 $i \in I_{\mu'}$ 。因为 $\mu' >_{\mu}$, 则有 $\mu'(i) >_i \mu(i) >_i i$, 所以 $\mu'(i) \in W$ 。令 $j = \mu'(i)$, 因为 μ, μ' 是稳定匹配, 那么在 μ, μ' 中是不会存在阻碍对 (i, j) 的, 即 $\mu'(j) >_{\mu} j$ 是不成立的。因此

$j \in W_\mu, \mu'(M_{\mu'}) \subset W_\mu$ 。

另一方面, 假设 $j \in W_\mu$, 同样的, 有 $\mu(j) > \mu'(j) > j$, 因此 $\mu(j) \in M$ 。令 $\mu(j) = i$, 因为 μ, μ' 是稳定匹配, 那么在 μ, μ' 中是不会存在阻碍对 (i, j) 的, 即 $\mu(i) > \mu'(i)$ 是不成立的。因此 $i \in M_{\mu'}, \mu(W_\mu) \subset M_{\mu'}$ 。

由于 μ' 和 μ 都是单射, 且 $M_{\mu'}$ 和 W_μ 都是有限集, 则可以得出, μ' 和 μ 与 $M_{\mu'}$ 和 W_μ 之间是一种双射关系。证毕。

定理2 假设 μ 是任意的一个稳定匹配, $S(\mu)$ 为与司机成功匹配的乘客的集合, $n(\mu)$ 为与司机成功匹配的乘客的数量。则有, 在所有的稳定匹配中, 成功匹配的乘客的集合 $S(\mu)$ 和数量 $n(\mu)$ 都是一致的。

证明: 采用反证法证明。假设偏好结构满足 $\gamma' = \gamma$, 司机 i 在稳定匹配 μ' 中是成功匹配的, 但是不在匹配 μ 中。则有 $i \in M_{\mu'}$, 根据定理1可知, $\mu(W_\mu) \subset M_{\mu'}$, 即 W_μ 通过 μ 与 $M_{\mu'}$ 形成一种映射关系, 那么司机 i 也必然在稳定匹配 μ 中, 这与假设矛盾, 因此命题成立。证毕。

推论1: 假设 μ 是通过 GS 算法司机 - optimal 求解的稳定匹配方案, (i, j) 是 μ 中的一个匹配对, 且 j 的偏好序列可以表示为: $\pi(j) = ([1], \dots, [i], [i] + 1, \dots, [k], [n_j], j)$, 那么将位置 $[i] + 1$ 到位置 $[n_j] + 1$ 之间的偏好信息从乘客 j 的偏好序中删除对匹配方案没有影响。同理, 假设 μ' 通过 GS 算法乘客 - optimal 求解的稳定匹配方案, 类似地对司机列表进行删减, 对整个匹配方案没有影响。

证明: 采用反证法证明。先证明通过司机 - optimal 的 GS 算法对乘客列表进行删减对匹配方案没有影响。令 μ 是通过 GS 算法司机 - optimal 求解的稳定匹配方案, (i, j) 是 μ 中的一个匹配对, 假设对乘客 j 的偏好序进行删减对匹配方案有影响, 通过定理1和定理2, 可以知道所有匹配中, 在方案 μ 中匹配成功的个体, 在其他稳定匹配方案中必然会匹配成功。那么说明在至少存在另一个稳定匹配方案 μ' , 使得 (k, j) 和 (i, l) 同时成立。因为在乘客 j 的偏好序中司机 k 的位置是在司机 i 后面的, 则有 $i > j, k$, 而由于是通过 GS 算法实现的司机 - optimal 的匹配方案, 那么 $u > u'$, 即 $j > l$, 由此定义3可知, (i, j) 将会是匹配方案 μ' 的一个阻碍对, 这与假设矛盾。因此命题成立。

同理, 通过乘客 - optimal 的 GS 算法对司机列表进行删减也不会对匹配方案产生影响。证毕。

3.1.2 Delete 算子构建

将 DA 算法应用到顺风车司机与乘客的匹配模型中, 司机 - Optimal 延迟接受算法 (DODA) 得到的匹配方案对司机来说是最优的, 但是对乘客来说却是最差的一种匹配方案。反之, 乘客 - Optimal 延迟接受算法 (RODA) 的流程, 得到的匹配方案对乘客来说是最优的。

(1) Delete 算子

DA 算法是 Delete 算子的核心组成部分, 利用 Delete 算子可以对匹配双方的偏好列表有效地进行精简, 记经过 Delete 算子精简之后的偏好序为 $\pi(p)_{gs}$ 。用一个列表来汇集所有司机的偏好序, 这里称这个列表为司机偏好列表, 记为 $pl(d)$, $pl(d) = [\pi(1); \pi(2); \dots; \pi(i); \dots; \pi(m)]$, $i \in I$ 。同理记乘客的偏好列表为 $pl(r)$, 则有 $pl(r) = [\pi(1); \pi(2); \dots; \pi(j); \dots; \pi(n)]$, $j \in J$ 。偏好序的精简过程如图2所示。第1步, 使用 MODA 算法后发现 d_3 与 r_5 成功匹配, 那么根据推论1可知, 将乘客 r_5 偏好序中位于 d_3 位置之后的元素 (即排序在5以后的元素) 删除, 对匹配结果没有影响。因为乘客 r_5 偏好序中位于 d_3 位置之后的元素都已经删除, 这意味着乘客 r_5 与 d_3 位置之后的司机都不可能进行匹配, 反之, 司机 d_2 和 d_6 也不可能与乘客 r_5 进行匹配, 即将 r_5 从司机 d_2 和 d_6 的偏好序中删除。依此类推对所有乘客和司机的偏好序进行初次精简。第2步, 使用 WODA 算法对司机偏好序进行精简, 如图2所示, 此时 (r_4, d_3) 是 WODA 算法求解的稳定匹配方案中的一个匹配对, 根据推论1可知, 对司机 d_3 偏好序中位于 r_4 位置之后的元素进行删除, 对匹配结果没有影响。与第一步精简方法同理, r_4 位置之后的元素 r_8, r_7, r_6 的偏好序中也可以进一步将 d_3 删除。

通过上述的两步基于 GS 算法的 Delete 算子, 形成一个精简的乘客偏好列表, 这里记为 $pl(r)_{gs} = [\pi(1)_{gs}, \pi(2)_{gs}, \dots, \pi(j)_{gs}, \dots, \pi(n)_{gs}]$, $j \in J$; 类似地, 记精简之后司机偏好列表为 $pl(d)_{gs} = [\pi(1)_{gs}, \pi(2)_{gs}, \dots, \pi(i)_{gs}, \dots, \pi(m)_{gs}]$, $i \in I$ 。令 M_{gs} 和 W_{gs} 分别表示通过 Delete 算子对司机和乘客的潜在可接受匹配对象进行删减之后的可行域。

3.1.3 算法流程

第1阶段的主要目的是求得司机与乘客集匹配双方的偏好列表。根据式(2)判断司机 i 与乘客 j 之间是否满足时间窗口约束, 如果满足时间窗口约束, 则根据式(6)和(7)分别计算 $u_i(i, j)$ 和

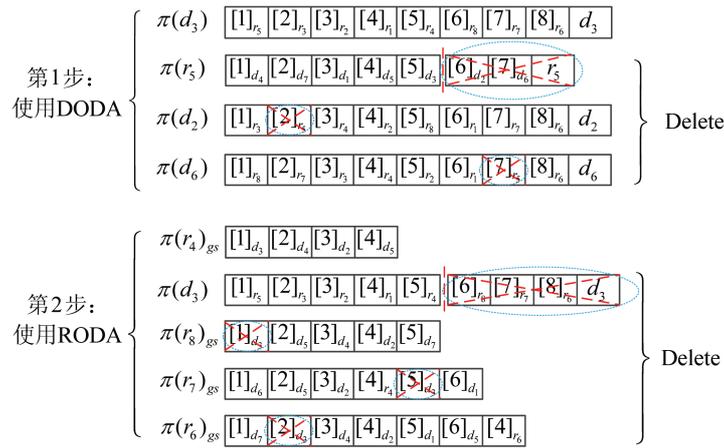


图2 Delete算子对偏好序精简过程

Fig. 2 Process of reduction of preference order by Delete operator

$u_j(i, j)$, 如果 $u(i, j) < 0$ 那么就是意味着, 尽管匹配对 (i, j) 满足时间窗口约束, 但是司机/乘客并不能通过匹配获得出行成本节省, 这种匹配是不稳定的, 只要某一方不能获得出行成本的节省, 即匹配不可行的, 此时令 $u_i(i, j) = u_j(i, j) = -K$, K 为一个足够大的正数。在求得所有司机与乘客的效用后, 进一步对司机与乘客的效用值进行排序, 这里采用最简单的冒泡法进行排序, 最终输出司机与乘客双方的偏好列表, 该偏好列表是按照效用值的大小从大到小依次排列的两个序列。

第2个阶段是对第1阶段得到的偏好列表进行精简, 最终输出精简的偏好列表。

3.2 模型处理

对参与者偏好列表进行精简之后, 模型 (11) 转化为:

$$\begin{aligned}
 f_{sta} &= \sum_{i \in M_{gs}} \sum_{j \in W_{gs}} x_{ij} (d_i - d_{ij}^o - d_{ij}^e) \\
 \text{s. t. } &\begin{cases} \sum_{i \in M_{gs}} x_{ij} \leq 1, \forall j \in J \\ \sum_{j \in W_{gs}} x_{ij} \leq 1, \forall i \in I \\ \sum_{i' \geq i} x_{i'j} + \sum_{j' \geq j} x_{ij'} + x_{ij} \geq 1, \forall i \in M_{gs}, j \in W_{gs} \\ x_{ij} \geq 0, \forall i \in M_{gs}, j \in W_{gs} \end{cases}
 \end{aligned}
 \tag{12}$$

定理3 模型 (12) 必然存在最优解^[20]。

证明: 本研究中参与者有两种配方式, 一是与可行的潜在匹配对象进行匹配或参与者与自己匹配, 那么集合 M_{gs} 和 W_{gs} 的大小都会远小于参与者的总数, 即小于 $m+n$ 。那么模型 (12) 为含有不超过 $m+n$ 个变量的规划模型, 故其最多具有 $2^{(m+n)^2}$ 个可行解, 即模型的可行解为有限多个。根据 GS 算法可知, 本

研究考虑的双边匹配问题必然存在稳定匹配解, 即模型 (12) 至少存在一个可行解。因此模型 (12) 必然存在最优解。

4 算例分析

顺风车的匹配成功与否往往与参与到顺风车平台的人数有关, 故将研究城区繁华区域的顺风车匹配, 由于实际的数据难以获得, 这里将采用合理的数据模拟方法来产生相应的出行供给与出行需求。

4.1 参数设置

假设城区繁华区域是一个长和宽都是 20 km 的一个正方形区域, 且所有出行车辆在该区域内的平均行驶速度都为 30 km/h, 其中, 乘客上下车的服务时间忽略不计, 为了研究参与者的出行模式的影响, 本研究将从研究两种不同的参与者起点与终点的分布入手: 一种是所有的参与者的起止点都均匀随机地分布在区域内, 另一种是参与者的起止点主要均匀随机地分布在 2 个中心。在后一种分布中, 司机乘客的起点和终点分别位于两个不同的中心内, 中心是一个圆形区域, 半径为 r 。

系统中参与者的时间窗口的产生方法如下: 为了提高顺风车的参与率, 主要研究上班高峰期的出行。根据已有研究可知, 出行者早上平均最晚出发时间为 7:30 am, 假设最晚出发时间是以 7:30 am 为均值, 标准差为 45 min 的一个正态分布。将均值转化为分钟, 则早上 7:30 可以记为 450 min, 即本研究的出行者最晚出发时间 ld_p 是以 450 为均衡, 45 为标准差的一个正态分布, 那么参与者 p 的最晚到达时间 $la_p = ld_p + t_p$, 其中 t_p 表示参与者 p 直接开车从

起点到终点的时间。

弹性时间 (Flexible Time): 是指除去参与者从起点出发到终点的行驶时间外, 可用的冗余时间。标记为 FT_p 。

$$FT_p = ld_p - ed_p = la_p - ea_p \quad (13)$$

假设弹性时间也是服从正态分布的, 均值为一个给定的值 FT_{min} , 标准差为 4 min, 那么根据式 (13) 可以求得 $ed_p = ld_p - FT_p$, $ea_p = la_p - FT_p$ 。

为了计算简便, 本研究假设车辆行驶单位距离的成本为固定的一个值 α , α 会随燃油价格的变化而变化, 每天不同的时间段 α 的取值也会发生变化, 如在滴滴顺风车服务中, 夜间单位行驶距离的收费明显比白天高出很多, 交通高峰期的收费也比非交通高峰期高, 但是本研究是一种静态的顺风车匹配方案, 故在某一个确定的时段, 这里的 α 的取值也将会是一个固定的值。假设顺风车匹配平台将会向每一个成功匹配的匹配对 (i, j) 收取一定的费用 φ_{ij} , 根据上文分析已知, 平台收取的费用与匹配对的距离节省成正比, 且比例记为 η , 为了计算方便, 令 η 是一个固定值。此外, 关于司机与乘客的时间机会成本 wt_p , 本研究假设时间机会成本的系数 ω 已知, 系数 ω 与参与者的工资水平成正相关, 假设参与者 p 的工资为 w 元/年, 假设每天工作时间为 8 h, 那么系数 ω 关系式可以表示为: $\omega_p = w / (12 \times 30 \times 8 \times 60)$ 。为了便于计算, 本研究假设所有参与者对时间机会成本的敏感程度是一致的, 以北京年平均工资为例计算时间机会成本系数, 根据国家统计年鉴, 北京市 2019 年平均年工资为 111 390 元, 那么 $\omega \approx 0.645$ 。

4.2 评价指标

(1) 匹配成功率 (Suc): 匹配成功率定义为成功匹配上的参与者出行量除以参与者发布的总的出行需求量。假设稳定匹配模型求解的结果中, 成功匹配的匹配对的数量为 N_s , 在平台中发布出行需求的司机数量为 N_d , 在平台中发布出行需求的乘客数量为 N_r , 那么匹配成功率为:

$$Suc = 2N_s / (N_d + N_r) \times 100\%$$

(2) 总行驶距离节省比率 (Sav): 总行驶距离节省比率定义为通过稳定匹配模型求解得到的系统总距离节省除以所有发布出行需求的司机与乘客直接从起点开车往终点的所有行驶距离之和, 可以用公式表示为:

$$Sav = f_{sta} / (\sum_{i \in N_d} d_i + \sum_{j \in N_r} d_j) \times 100\%$$

(3) 参与者个人出行平均节省比率 (Si): 参与者个人出行平均节省定义为参与者出行的个人距离节省除以其独自出行的行驶距离, 再对所有成功匹配的参与者求平均值。根据不同的成本分配方式, Si 的表达式是不一样的, 本研究主要讨论按比例分配方式。假设参与者按比例分配距离节省, 那么 Si 的公式表达为:

$$Sipr = \frac{(\sum_{i,j} x_{ij} \frac{d_i}{d_i + d_j} \frac{d_i - d_{ij}^o - d_{ij}^e}{d_i} + \sum_{i,j} x_{ij} \frac{d_j}{d_i + d_j} \frac{d_i - d_{ij}^o - d_{ij}^e}{d_j})}{2N_s} \times 100\%$$

(4) 司机平均绕路比率 (Dt): 司机平均绕路比率定义为司机因为参与顺风车出行需要的额外行驶距离除以该司机单独出行的行驶距离, 再对所有成功匹配的司机求取平均值。 Dt 的表达式如下:

$$Dt = (\sum_{i,j} x_{ij} \frac{d_j + d_{ij}^o + d_{ij}^e - d_i}{d_i}) / N_s \times 100\%$$

为了衡量本研究构建的稳定匹配模型以及提出的 Delete 算子的稳定性, 本节进一步构建了如下指标来进行衡量模型的效果。

(5) Price of Anarchy (POA): 根据前文定义, POA 定义为稳定距离总节省偏离系统最优距离总节省的比例。 POA 表达式如下:

$$POA = (f_{so} - f_{sta}) / f_{so} \times 100\%$$

(6) 程序运行时间: 本研究可以分为 3 种: 直接求解系统最优值 f 已经对应的匹配方案的运行时间, 记为 RT ; 求解考虑参与者之间成本分配的系统最优值 f_{so} 的运行时间, 记为 RT_{so} ; 利用 Delete 算子对参与者偏好列表精简之后, 求解稳定匹配值及对应的匹配方案的运行时间, 记为 RT_{sta} 。

4.3 结果分析

(1) 参与者数量、聚集中心对匹配的影响

将 Suc , Sav , $Sipr$, Dt , POA 5 个指标对平台中潜在参与者数量和匹配区域中聚集中心对顺风车匹配结果的影响进行对比分析。如图 4 所示, 参与者数量变化区间为 400 ~ 2 400, 参与者弹性时间的均值为 $FT = 30$ min, 弹性时间服从正态分布, 标准差为 4 min, 平台向成功匹配的匹配对收取的费用比例 $\eta = 10\%$, 这里并不考虑参与者的时间机会成本, 即 $\omega = 0$, 区域内所有司机单位行驶距离成本 α 为 2 元/km。

图 3 (a) 是参与者数量与稳定匹配率关系的对比图, 从图中可以看到, 有交通中心时的稳定匹配

率远远高于没有交通中心时,当服务区内参与者数量仅为400时,有交通中心的顺风车稳定匹配率就已经达到了91.3%,此时,无交通中心的顺风车稳定匹配率仅仅为62.6%;而当服务区内参与者数量上升至2400时,有交通中心的顺风车稳定匹配率达到了95%,无交通中心的顺风车匹配率仅为78%。图3(b)是参与者数量与稳定匹配总距离节省比率关系的对比图,从图中可以看到,有交通中心时的距离节省比率也是远远高于没有交通中心时,从35.1%上升到39.9%,无交通中心的稳定匹配总距离节省比率随着参与者数量的变化从18.3%变动到27.2%。图3(c)是参与者数量与稳定匹配个人平均出行距离节省比率关系的对比图,与图3(a)和图3(b)中的曲线有着类似的规律,随着参与者人数的不断增加, $Sipr$ 曲线不断上升,且有两个交通中

心的个人平均距离节省比率远远高于没有交通中心。图3(d)是参与者数量与司机平均绕路比率关系的对比图,从图中可以看出,随着参与者数量的不断增加,司机的平均绕路比率不断下降,且有交通中心时司机的绕路比率是远低于没有交通中心时。当参与者人数上升至2400时,在有两个交通中心的情况下,司机的绕路比率仅仅为16.1%,而此时,在没有交通中心的情况下,高比率为24.2%。此外需要指出的是,尽管在有两个交通中心的情况下,顺风车匹配的指标 Suc , Sav , $Sipr$, Dt 都展示出更好的指标值并一直保持着良好的稳定性能。Gap 指标,也就是 POA 值,一直维持在6.4%到7.4%之间。由此可见,随着参与者人数的增加,本研究提出的稳定优化模型及算法都展示了良好,得到的稳定匹配解非常接近系统最优解。

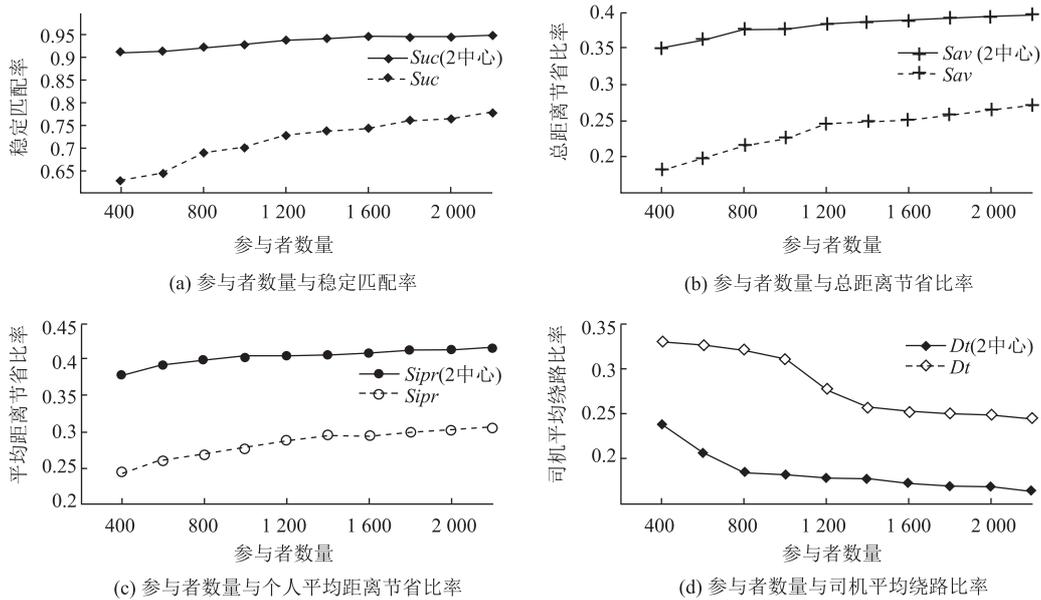


图3 参与者数量、聚集中心对顺风车匹配的影响

Fig. 3 Influence of number of participants and gathering center on ridesharing matching

(2) 时间机会成本、聚集中心对匹配的影响

下面将对参与者时间机会成本和匹配区域中聚集中心对顺风车匹配结果的影响进行对比分析。其中参与者时间机会成本系数从0~0.9变化。参与者弹性时间的均值为 $FT = 30$ min,弹性时间服从正态分布,标准差为4 min,平台向成功匹配的匹配对收取的费用比例 $\eta = 10\%$,区域内参与者人数为1400,其中司机700,乘客700,区域内所有司机单位行驶距离成本 α 为2元/km。

图4(a)是参与者时间机会成本系数与稳定匹配率关系的对比图,从图中可以看到,无论服务区内有无聚集中心,稳定匹配率总是随着参与者时间

机会成本系数的变大而不断下降,不同的是有两个聚集中心的情况下,稳定匹配率处于较高的状态。当参与者的时间机会成本系数增大到0.9时,匹配率依然高达77.6%,而此时没有聚集中心时的稳定匹配率仅仅为48.6%,匹配率较低。图4(b)是参与者时间机会成本与稳定匹配总距离节省比率关系的对比图,与图(a)中的曲线有着类似的规律,随着参与者时间机会成本系数的不断增大, Sav 曲线不断下降。图4(c)是参与者时间机会成本与稳定匹配个人平均出行距离节省比率关系的对比图,从图中可以发现, $Sipr$ 曲线不但没有随着时间机会成本系数的增加而下降,反而一直保持比较稳定的比率,

并呈现略微的上升趋势, 原因在于出行者对时间的敏感程度高, 往往不接受需要花太多时间的匹配, 即拒绝绕路距离较长的匹配, 接受的大都是绕路时间短, 且成本节省比率大的匹配, 因而会使个人出行的平均距离节省比率呈现上升趋势。而这一点与图 4 (d) 中参与者时间机会成本与司机平均绕路比率关系是相呼应的, 从图中可以看出, 随着参与者时间机会成本系数的不断增加, 司机的平均绕路比

率是不断下降的。图 4 (e) 是参与者时间机会成本与稳定匹配 POA 值关系的对比图, 如图所示, 随着时间机会成本系数的增大, 稳定匹配解与系统最优解之间的差距越来越大, 原因如前所述, 出行者时间敏感度高时, 出行者很有可能会拒绝一些耗时有而距离节省的匹配, 导致稳定匹配解与最优解之间的差距越来越大。但是需要指出的, 对时间很敏感的人群参与到顺风车出行当中的比率相对较小。

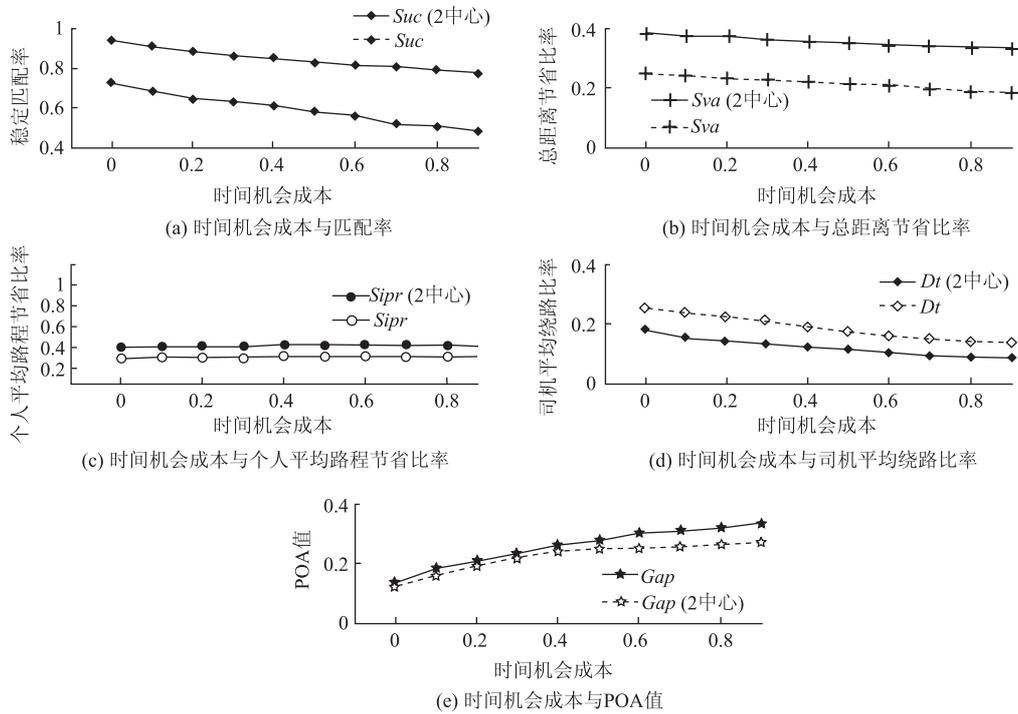


图 4 弹性时间、聚集中心对顺风车匹配的影响

Fig. 4 Influence of elastic time and gathering center on ridesharing matching

(3) 算法效率

这里将对 3 种不同的求解方式的程序运行时间进行对比: 第 1 种, 假设参与者之间按原始出行距离的比例分配出行成本, 可以保证得到的匹配方案中所有的成功匹配的司乘双方都有成本节省, 对应模型为式 (10), 记对应的程序运行时间为 RT_{so} ; 第 2 种, 为系统稳定匹配模型 (11) 对应的程序运行时间 RT_{sta2} , 此时参与者的偏好列表并未进行精简; 第 3 种, 利用 Delete 算子对参与者偏好列表精简之后模型 (12), 求解稳定匹配值及对应的匹配方案的运行时间, 记为 RT_{sta1} 。

对 RT_{sta1} , RT_{sta2} , RT_{so} 的比较可见图 5, 如图中所示, 随着参与者规模的不断增大, RT_{sta1} , RT_{sta2} , RT_{so} 的值都有着明显的增加, 在参与者规模小于 1 400 时, 求解模型 (10) 的程序运行时间 RT_{so} 、求解没有精简偏好列表的模型 (11) 对应的程序运行

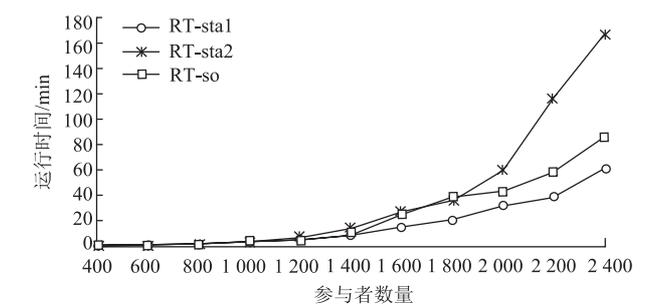


图 5 各程序运行时间对比

Fig. 5 Comparison of running time of each program

时间 RT_{sta2} 和使用精简列表的模型 (12) 的程序运行时间 RT_{sta1} 差距很小, 尤其是在小规模问题时, 往往出现 $RT_{sta2} < RT_{sta1}$ 和 $RT_{so} < RT_{sta1}$ 的情况, 说明本研究提出的算法对解决小规模顺风车匹配并未展现出优势, 当参与者规模大于 1 400 后, RT_{sta1} 曲线的增长趋势相对而言是相当缓慢的; 当参与者规模达

到 2 400 时, RT_{sta2} 的值将近是 RT_{sta1} 值的 3.24 倍, 而 RT_{so} 的值将近是 RT_{sta1} 值的 1.68 倍。这意味着, 本研究提出的 Delete 算子能够有效地对偏好列表进行精简, 且精简之后使得模型的求解速度大幅度增加, 说明该算法能够有效解决大规模的顺风车匹配问题。

5 结论

本研究构建了顺风车稳定匹配优化模型, 并从参与者数量、参与者弹性时间、参与者时间机会成本、区域内有没有聚集中心和算法效率等 5 个方面对本研究提出的算法进行分析, 其中衡量指标包含匹配率 Suc 、总距离节省比率 Sav 、参与者个人出行距离节省比率 $Sipr$ 、司机绕路比率 Dt 以及稳定优化解和系统最优解之间的 POA 值。通过全面的试验和结果分析发现: (1) 无论服务区内有无聚集中心, 随着参与者数量的增加, 匹配率 Suc 、总距离节省比率 Sav 和参与者个人出行距离节省比率 $Sipr$ 都不断上升, 但服务区内有两个交通聚集中心时, Suc , Sav 和 $Sipr$ 都分别高于没有交通中心时, 而随着参与者数量的增加, 司机绕路比率 Dt 不断下降; (2) 无论服务区内有无交通聚集中心, 随着参与者时间机会成本系数的增加, 匹配率 Suc 、总距离节省比率 Sav 和司机绕路比率 Dt 是不断下降的, 而随着参与者时间机会成本系数的增加, POA 值不断增加的; (3) 随着参与者弹性时间的增加, 对于服务区内无交通中心的情况, 随着参与者数量的增加, 匹配率 Suc 、总距离节省比率 Sav 和参与者个人出行距离节省比率 $Sipr$ 都不断上升, 司机绕路比率 Dt 是不断下降的; (4) 通过比较系统最优化模型、考虑参与者成本的系统优化模型和稳定匹配优化模型求解时间, 发现本研究提出的偏好列表精简算法非常有效, 能够有效地降低求解稳定匹配解的时间, 使得本研究提出的稳定匹配优化模型能够应用到大型的顺风车匹配问题中。总而言之, 本研究提出算法能够快速有效地解决单司机单乘客稳定匹配问题。

参考文献:

References:

- [1] ZHOU G, HUANG K, MAO L. Design of Commute Carpooling Based on Fixed Time and Routes [J]. International Journal of Vehicular Technology, 2014 (1): 1 - 9.
- [2] AGATZ N A H, ERERA A L, SAVELSBERGH M W P, et al. Dynamic Ride-sharing: A Simulation Study in Metro Atlanta [J]. Transportation Research Part B: Methodological, 2011, 45 (9): 1450 - 1464.
- [3] MA S, WOLFSON O. Analysis and Evaluation of the Slugging Form of Ridesharing [C] // Proceedings of the ACM International Symposium on Advances in Geographic Information Systems. Orlando: ACM, 2013: 64 - 73.
- [4] RODIER C, ALEMI F, SMITH D. Dynamic Ridesharing: Exploration of the Potential for Reduction in Vehicle Miles Traveled [J]. Transportation Research Record, 2016, 2542: 120 - 126.
- [5] HERBAWI W, WEBER M. The Ridematching Problem with Time Windows in Dynamic Ridesharing: A Model and a Genetic Algorithm [C] // 2012 IEEE Congress on Evolutionary Computation (CEC). Brisbane: IEEE, 2012: 1 - 8.
- [6] 邵增珍, 王洪国, 刘弘, 等. 多车辆合乘问题的两阶段聚类启发式优化算法 [J]. 计算机研究与发展, 2013 (11): 2325 - 2335.
- SHAO Zeng-zhen, WANG Hong-guo, LIU Hong, et al. Heuristic Optimization Algorithms of Multi-carpooling Problem Based on Two Stage Clustering [J]. Journal of Computer Research and Development, 2013 (11): 2325 - 2335.
- [7] WINTER S, NITTEL S. Ad Hoc Shared-ride Trip Planning by Mobile Geosensor Networks [J]. International Journal of Geographical Information Science, 2006, 20 (8): 899 - 916.
- [8] STIGLIC M, AGATZ N, SAVELSBERGH M, et al. The Benefits of Meeting Points in Ride-sharing Systems [J]. Transportation Research Part B: Methodological, 2015, 82: 36 - 53.
- [9] ARSLAN A M, AGATZ N, KROON L, et al. Crowdsourced Delivery: A Dynamic Pickup and Delivery Problem with Ad Hoc Drivers [J]. Transportation Science, 2019, 53 (1): 222 - 235.
- [10] HERBAWI W, WEBER M. Evolutionary Multiobjective Route Planning in Dynamic Multi-hop Ridesharing [C] // European Conference on Evolutionary Computation in Combinatorial Optimization. Berlin: Springer, 2011: 84 - 95.
- [11] CALVO R W, DE LUIGI F, HAASTRUP P, et al. A Distributed Geographic Information System for the Daily Car Pooling Problem [J]. Computers & Operations Research, 2004, 31 (13): 2263 - 2278.
- [12] TEODOROVIC D, DELL'ORCO M. Bee Colony Optimization: A Cooperative Learning Approach to Complex Transportation Problems [J]. Advanced OR and

- AI methods in Transportation, 2005 (1): 51 – 60.
- [13] GALE D, SHAPLEY L S. College Admissions and the Stability of Marriage [J]. American Mathematical Monthly, 1962, 69 (1): 9 – 15.
- [14] IWAMA K, MIYAZAKI S. A Survey of the Stable Marriage Problem and Its Variants [C] // International Conference on Informatics Education and Research for Knowledge-circulating Society. Kyoto: IEEE, 2008: 131 – 136.
- [15] GEISBERGER R, LUXEN D, NEUBAUER S, et al. 10th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'10). Dagstuhl, Germany: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2010: 88 – 99
- [16] KLEINER A, NEBEL B, ZIPARO V A. A Mechanism for Dynamic Ride Sharing Based on Parallel Auctions [C] // Proceedings of the Twenty-second International Joint Conference on Artificial Intelligence. Palo alto: AAAI Press, 2011: 266 – 272.
- [17] WANG X, AGATZ N, ERERA A. Stable Matching for Dynamic Ride-sharing Systems [J]. Transportation Science, 2018, 52 (4): 850 – 867.
- [18] VATE J H V. Linear Programming Brings Marital Bliss [J]. Operations Research Letters, 1989, 8 (3): 147 – 153.
- [19] ROTH A E, ROTHBLUM U G, VANDE VATE J H. Stable Matchings, Optimal Assignments, and Linear Programming [J]. Mathematics of Operations Research, 1993, 18 (4): 803 – 828.
- [20] 樊治平, 李铭洋, 乐琦. 考虑稳定匹配条件的双边满意匹配决策方法 [J]. 中国管理科学, 2014 (4): 112 – 118.
- FAN Zhi-ping, LI Ming-yang, YUE Qi. Decision Analysis Method for Two-sided Satisfied Matching Considering Stable Matching Condition [J]. Chinese Journal of Management Science, 2014 (4): 112 – 118.

(上接第 102 页)

- CHEN Wei-zhong, NI Zong-zan, PAN Xiao-ping, et al. Receiver Operating Characteristic Curves to Determine the Optimal Operating Point and Doubtable Value Interval [J]. Modern Preventive Medicine, 2005, 32 (7): 729 – 731.
- [16] 高惠璇. 处理多元线性回归中自变量共线性的几种方法: SAS/STAT 软件 (6.12) 中 REG 等过程增强功能的使用 [J]. 数理统计与管理, 2016, 19 (5): 49 – 55.
- GAO Hui-xuan. Some Method on Treating the Collinearity of Independent Variables in Multiple Linear Regression [J]. Journal of Applied Statistics and Management, 2016, 19 (5): 49 – 55.
- [17] 归庆明, 姚绍文, 顾勇为, 等. 诊断复共线性的条件指标 – 方差分解比法 [J]. 测绘学报, 2006, 35 (3): 210 – 214.
- GUI Qing-ming, YAO Shao-wen, GU Yong-wei, et al. A New Method to Diagnose Multicollinearity Based on Condition Index and Variance Decomposition Proportion (CIVDP [J]. Acta Geodaetica et Cartographica Sinica, 2006, 35 (3): 210 – 214.
- [18] 吴喜之. 应用回归及分类: 基于 R [M]. 北京: 中国人民大学出版社, 2016.
- WU Xi-zhi. Applied Regression and Classification: Based on R [M]. Beijing: China Renmin University Press, 2016.
- [19] 王济川, 郭志刚. Logistic 回归模型: 方法与应用 [M]. 北京: 高等教育出版社, 2001.
- WANG Ji-chuan, GUO Zhi-gang. Logistic Regression Models: Methods and Application [M]. Beijing: Higher Education Press, 2001.