

大数据背景下动态共乘的研究进展

沈弼龙¹ 赵颖¹ 黄艳² 郑伟民¹

¹(清华大学计算机科学与技术系 北京 100084)

²(北德克萨斯州大学计算机科学与工程系 美国德克萨斯州登顿市 311277)

(shenbilong@gmail.com)

Survey on Dynamic Ride Sharing in Big Data Era

Shen Bilong¹, Zhao Ying¹, Huang Yan², and Zheng Weimin¹

¹(*Department of Computer Science and Technology, Tsinghua University, Beijing 100084*)

²(*Computer Science and Engineering Department, University of North Texas, Denton, TX, USA 311277*)

Abstract The availability of multi-source data in big data era can potentially lead to a revolution in ride sharing, which has been widely studied in academia as a means of reducing the number of cars, congestion, and pollution by sharing empty seats, and is lack of popularity in practice due to the inflexibility of off-line booking, limited resources and so on. In big data era, dynamic ride sharing, powered by mobile computation, location based service, and social networks, emerges and gains popularities recently for providing real-time and flexible ride sharing services through real-time travel planning systems. These systems raise research opportunities as well as challenges, including how to process real-time location data and traffic data, to match ride requests and cars in real time, and to provide fair, secure, and low-priced services to gain more participants. This paper defines the problem of dynamic ride sharing formally and discusses its variants and recent developments. The framework of filter and refine to solve the real-time challenges of matching requests and cars is then discussed. In particular, in the filter step, we introduce the method of pre-computing, spatio-temporal index, grouping, and parallelizing. In the refine step, we introduce the method of lazy calculation strategy, new index tree structure, and evolutionary computation. We also discuss the techniques related to social factors such as pricing strategies, credit system, human-computer interaction, and security in big data era. Finally, this paper ends with a panoramic summary and a discussion on possible future research directions.

Key words ride sharing; dynamic ride sharing; big data; optimization algorithm; urban computing

摘要 共乘也被称为“合乘”、“拼车”、“顺风车”，通过有效整合运力资源减少路上行驶车辆数量，对缓解交通拥堵、降低出行费用、减轻环境污染都有重要意义。大数据背景下实时更新的车辆位置信息数据、城市交通数据、社交网络数据，为智能出行特别是共乘带来了全新的发展机遇。在车辆行驶中对乘客请求进行实时匹配的动态共乘，是大数据背景下智能出行发展趋势的代表。在统一归纳了解决动态共乘实时性的 Filter and Refine 框架基础上，介绍了动态共乘的各种类型；针对大数据背景下动态共乘问题遇到的问题，对 Filter 步骤中预先计算可行解、建立动态空间索引、基于请求分组预处理及并行优化方法，Refine 步骤中简化计算模型、采用新型数据结构、利用启发式算法等优化方法进行了详细介绍；然后对大数据背景下保证动态共乘系统的价格机制、信用体系和人机接口等相关技术进行了分析；最后，总结展望了大数据背景下动态共乘中亟待解决的关键问题和未来的研究方向，以期创造低碳生活、绿色出行，解决环境污染有所启示。

关键词 共乘;动态共乘;大数据;优化算法;城市计算

中图分类号 TP391

随着经济发展、人民生活水平提高,私家车的普及率也越来越高,给生活带来了方便,但车辆数量的增加也带来了交通拥堵、出行费用上涨、环境污染等一系列问题。根据中国社会科学院数据显示,2012年北京因交通拥堵每天增加社会成本4 000万元,每年经济损失达146亿元;同时机动车尾气也是雾霾^[1]的主要成因之一。共乘出行中驾驶员把车辆上空余的位置提供给乘客,也被称为“拼车”、“汽车共享”、“顺风车”或“搭便车”。这种出行方式提高了交通资源的利用率,既降低了出行成本,又缓解了交通拥堵、雾霾、噪音等环境污染问题^[2-4]。据研究在美国只需4%的司机选择共乘,就可以解决美国68个主要城市次年车辆增长造成的拥堵问题^[5]。早期共乘以上下班通勤和长途旅行为主,甚至传统的公交车、地铁在某种程度上都可以看成更为广义的共乘系统。共乘方式按出发后行程是否可变,分为静态共乘和动态共乘2种。静态共乘需在车辆出发前规划好行程,出行之后不能改变行程,这种共乘方式灵活性较差,没有发挥资源最大潜力,对共乘服务的大规模发展有一定限制。随着网络基础设施的不断完善,智能手机等移动终端的广泛应用及GPS、北斗等定位技术的发展,为实时动态共乘发展提供了数据服务基础。动态共乘允许司机在车辆行驶中与乘客匹配,使得共乘匹配更为灵活、方便,但同时带来了计算复杂度、匹配成功率、价格机制、信用机制等一系列问题。本文的贡献主要在于:1)对动态共乘问题进行了统一框架的定义;2)多维度介绍了共乘问题分类;3)对动态共乘问题中遇到的实时性等挑战和解决方案进行了分析总结;4)对影响共乘问题的价格机制、信用体系和人机接口的技术因素进行了分析和总结;5)通过对当前动态共乘问题及典型应用研究分析,指出了动态共乘面临的主要问题和未来发展趋势。

1 共乘问题概述

1.1 共乘发展历史

共乘源自于熟人之间的自发行为,这种行为可以直接减少用户出行成本。以北美为例^[6-7],共乘发展主要经历了5个阶段:1)1942年共乘以汽车俱乐部的形式出现,主要是为了解决二战期间的资源短缺问题,在这一阶段处于共乘萌芽阶段,主要依靠简

单的公告牌来完成;2)在20世纪60年代到70年代之间,在这阶段出现了工人自发和政府组织的共乘行为,如美国在一些高速公路上修建专门的共乘车道^[8](high-occupancy vehicle lane, HOV),规定只有上座2人以上的车才可以驶入,同时Slugging^[9]形式的长途拼车也出现了,共乘的人员范围进一步扩大;3)从20世纪80年代开始,以解决环境污染和交通拥堵为目标,运输管理协会等组织出现,基于电话叫车服务的共乘业务开始发展;4)从1999年到2004年,围绕解决拥堵问题,共乘服务商开始大力发展参与共乘的人数,基于互联网的在线共乘业务和专门的出行服务511^[10]也在此阶段出现,智能共乘开始兴起;5)从2004年到现在,随着智能手机等移动终端的普及,GPS、北斗等定位技术的发展,海量的移动终端与车辆位置信息可以方便被获取,基于移动终端的共乘平台开始涌现,共乘进入了全新以技术为驱动的时代,灵活的实时动态共乘相关研究与原型系统开始出现。新的技术为共乘解决气候变暖、降低石油进口和解决交通拥堵等问题提供了更有力的保证。欧洲国家^[11-12]和日本、新加坡等国也都开始了共乘的尝试。

1.2 共乘发展现状

随着定位技术、移动网络的发展和智能终端的普及,终端设备、即时通信、城市交通、社交网络每天都在产生海量数据,而且这些数据量也在随着时间迅速增长,以江苏省为例,其交通轨迹数据已高达500多亿条、100多TB,每日新增数据量7 000万条、150多GB^[13]。以海量个人的实时出行数据、交通信息数据、社交媒体数据等为代表的大数据背景为智能交通系统发展带来了新的机遇^[14]。共乘这种出行方式出现已久,但由于过去信息传播能力与手段的限制,很长一段时间仅限于具有亲密关系的人群,没有大规模发展。而在大数据背景下,实时地理位置感知、海量数据处理等技术的成熟为共乘的大规模发展提供了平台支撑,基于终端设备的出行方式被人们逐步接受为共乘的大规模发展提供了用户支撑,共乘出行迎来了全新的发展机遇。

近年来智能共乘平台逐步普及发展。国外已涌现出Avego^[15], mitfahrgelegenheit^[16], zimride^[17], Carpooling^[18], Blablacar^[19]等共乘平台, Uber^[20], Lyft^[21]用车平台也推出拼车服务。国内的共乘业务

相对起步较晚,在杭州、北京、武汉等城市,相继出现了民间自发组织的共乘,如武汉的邻里合乘^[22]项目,北京 2008 年因奥运期间限行而自发组织的拼车等.近年随着滴滴^[23]、神州专车^[24]等智能终端叫车业务的发展,国内用户对通过智能终端叫车的出行方式也已经接受.共乘业务服务商也如雨后天春笋般涌现,已有 51 用车^[25]、嘀嗒拼车^[26]、滴滴顺风车^[23]等数十个拼车平台,并在上下班通勤中得到了大量应用.

从现有共乘应用形式上看,车主与乘客都是在行程出发之前完成匹配,匹配后按预先规划路线行驶.这种共乘对资源的利用率有一定的局限性,车辆如果出发之前没有完成乘客匹配,在行程出发后即使有可匹配的乘客也无法对其请求进行响应,实载率提高有限,在应用上存在一定局限性,没有充分发掘交通大数据的潜力.为进一步提高共乘资源利用率,近年研究学者提出了动态共乘的概念^[27],并有大量研究工作,在提高匹配度^[28-31]、降低行程开销^[4,28-29,32-38]、减少用户请求响应时间^[4,29-30,32,34-40]等方面展开了研究,其中许多研究基于真实出行数据集进行了测试仿真.例如黄艳^[37],郑宇^[4],Agatz^[28]等人分别利用了上海市 17 000 台出租车行程轨迹数据^[37]、北京市 33 000 台出租车轨迹数据、亚特兰大市的真实出行模型进行了共乘出行的模拟.实验模拟表明,动态共乘能有效降低车辆空车率,提高实载率,为节省出行成本、降低碳排放、缓解交通拥堵

提供了有效的解决方案,是未来智能出行的发展方向之一.

2 动态共乘问题定义

共乘问题可看成优化问题,即对路上行驶的 n 辆车、 m 个服务请求,规划出车辆与服务请求的匹配关系及行车路线,使得到满足的服务请求数最多.共乘按车辆出发后能否进行匹配,分为静态共乘和动态共乘 2 类.静态共乘,司机和乘客需在行程开始前,完成满足乘客和路程约束条件的匹配,出发后不能更改行程;动态共乘,司机在行程开始之后,仍可在满足乘客和路程约束的情况下,与乘客进行实时动态匹配.静态共乘可以看成是动态共乘在某一时刻收到所有请求的特殊情况.动态共乘依赖网络、智能终端、精准定位技术,采用动态匹配算法,车辆可在行驶中接受新的乘客请求,进行实时动态共乘匹配,使运力资源得到了最大化的利用.动态共乘带来了匹配方法、路线规划、法律规范、出行安全、个人隐私等问题,本文中主要对大数据背景下动态共乘中的匹配方法与路线规划问题及促进动态共乘的相关技术进行讨论.

2.1 动态共乘系统概览

如图 1 所示,一个典型的动态共乘请求处理流程可以表示为:1)乘客将行程请求及相应约束条件发送到服务器;2)服务器根据乘客提供的行程请求及相应约束,将乘客与能提供相应服务的车辆进行

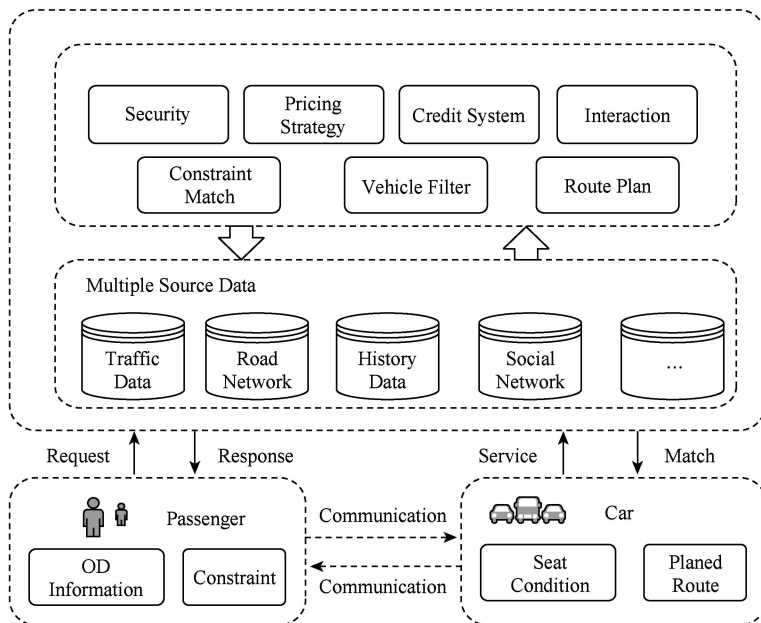


Fig. 1 Dynamic ride sharing system framework

图 1 动态共乘系统框架示意图

匹配;3)对符合乘客要求的车辆,将匹配结果发送给相应司机,若匹配失败直接通知乘客请求失败;4)司机根据收到的用户请求结合自身情况进行服务确认;5)服务器将匹配好的服务信息通知给乘客;6)司机按新的行程安排将乘客载上车并送到指定地点。

2.2 动态共乘基本概念

定义 1. 路网. 路网为集合 $G = \langle V, E, W \rangle$. 其中, V 为节点集合, 表示路网中不同的位置点; E 为路网中各边 $(u, v) \in E, (u, v \in V)$ 的集合; 每条边都有自己的权值, 记为 $W(u, v)$, 代表通过边 (u, v) 的开销, 开销用时间或距离表示. 给定行驶速度下, 距离和时间开销度量可相互转换。

定义 2. 服务请求. 给定 $G = \langle V, E, W \rangle$, 服务请求 $tr = \langle o, d, p_n, t_{ow}, t_{dw}, \epsilon \rangle$. 其中, 请求的起点 $o \in V$; 终点 $d \in V$; p_n 为本次请求的乘客人数; t_{ow} 为上车时间窗口, $t_{ow} = \langle t_{o,s}, t_{o,e} \rangle$, $t_{o,s}$ 为上车的最早时间, $t_{o,e}$ 为上车的最晚时间; t_{dw} 为下车时间窗口, $t_{dw} = \langle t_{d,s}, t_{d,e} \rangle$, $t_{d,s}$ 为下车最早时间, $t_{d,e}$ 为下车最晚时间, 如图 2 所示; 服务约束 ϵ 用来表示最大容许的绕路开销, 为 $(1 + \epsilon) \times dist(o, d)$, 其中 $dist(o, d)$ 表示为点 o 与点 d 之间的距离. 定义车辆实际到达 $tr.o$ 的时间为 $tr.t_{vo}$, 将乘客送达 $tr.d$ 的时间定义为 $tr.t_{vd}$.

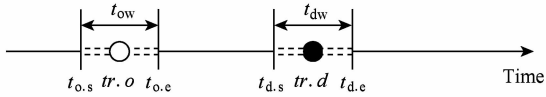


Fig. 2 Request time window

图 2 服务请求时间窗口示意图

将给定路网 G 上, 当前时刻所有行程请求的集合记为 TR .

定义 3. 车辆. $Car = \langle id, t, loc, schedule, TR_{rec}, n_{cap}, n_{emp} \rangle$. 其中, id 表示车辆唯一编码; t 表示时间戳; loc 表示车辆在时刻 t 的位置; $schedule$ 表示车辆时刻 t 已有安排; TR_{rec} 表示当前车上已有乘客请求集合, $TR_{rec} = \langle tr_1, tr_2, tr_3, \dots \rangle$; n_{cap} 表示车辆总容量; n_{emp} 表示车上当前空闲容量。

将给定路网 G 上, 当前时刻所有车辆的集合记为 $C = \langle Car \rangle$.

定义 4. 行程安排. 给定路网 G , 行程安排是某一时刻车辆 Car 对其将要完成上客点、落客点的时序安排, $schedule = \langle v_0, v_1, \dots, v_k \rangle$ 这段行程可以表示为一个顶点序列 (v_0, v_1, \dots, v_k) , 其中 (v_i, v_{i+1}) 为集合 E 中的边, v_0 表示行程安排时车辆的位置点。

定义 5. 动态共乘问题. 给定路网 G 上的车辆

集合 C 和服务请求集合 TR , 将不同的服务请求分配到不同车辆并进行行程安排, 使车辆在满足动态共乘约束的条件下, 达成一定的优化目标。

在后面的 2.3 节和 2.4 节中将分别介绍动态共乘需满足的约束条件和优化目标。

2.3 动态共乘约束条件分析

定义 6. 有效分配. 给定车辆 Car 与一个单独的服务请求 tr , 当他们的匹配满足共乘约束时, 称该匹配为有效分配。

定义 7. 请求完成. 一个单独的服务请求 tr 被分配到车辆 Car , 在满足共乘约束的条件下, 该车将乘客由 $tr.o$ 送至 $tr.d$ 处, 则称请求 tr 被 Car 完成。

动态共乘主要有时间、距离、价格及个人喜好等约束条件. 时间约束有服务时间窗口、到达时间点和额外增加时间等; 距离约束条件有整体行程距离、个人绕路距离等; 价格约束有成本开销、最低成交价格等; 个人喜好约束有个人性别、行为习惯、驾驶经验等. 个人喜好约束属人文范畴, 不在本文研究范围内. 时间和距离如前所述在给定行驶速度下, 可以互相转换; 价格可通过定价策略基于行程和时间获得, 因此本文主要以路网开销 $W(u, v)$ 及距离作为分析基础。

定义 8. 行程开销. 对于一个车辆 Car 的行程安排, 完成 $schedule = \langle v_0, v_1, \dots, v_k \rangle$ 中各点的行程, 这段行程的总开销定义为

$$W(schedule) = \sum_{i=0}^{k-1} W(v_i, v_{i+1}).$$

定义 9. 绕路开销. 某个行程中的车辆 Car 在没有新乘客加入时路程开销为 W_i , 新乘客加入后其路程开销为 W_j , 定义由新增乘客而给原来行程带来的绕路开销距离 $W\Delta_{detour}(v_i, v_j) = W_j - W_i$, 绕路开销

$$\delta_{det} = \frac{W\Delta_{detour}(v_i, v_j)}{W_i}.$$

如图 3 所示, 新增乘客而给原来行程带来的绕路开销比:

$$\delta_{det} = \frac{W(v_1, v_3) + W(v_3, v_4) + W(v_4, v_2) - W(v_1, v_2)}{W(v_1, v_2)}.$$

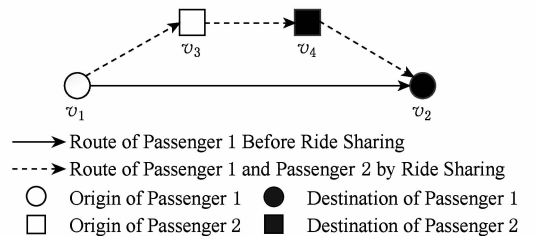


Fig. 3 Detour cost

图 3 绕路开销示意图

在本文中,对某台车辆 Car 和某个乘客请求 tr_i ,定义动态共乘有效分配的约束条件如下:

$$tr_i.p_n \leq Car.n_{emp}, \quad (1)$$

乘客可在规定时间窗口内完成请求:

$$\forall tr_i \in Car.TR_{rec},$$

$$\begin{aligned} tr_i.t_{o,s} &\leq tr_i.t_{vp} \leq tr_i.t_{o,e}, \\ tr_i.t_{d,s} &\leq tr_i.t_{vd} \leq tr_i.t_{d,e}, \end{aligned} \quad (2)$$

对每个请求,乘客绕路开销比小于最大容忍值:

$$\delta_{det} \leq \epsilon. \quad (3)$$

近年来,不同动态共乘研究中的约束条件如表 1 所示:

Table 1 Constraint Conditions of Dynamic Ride Sharing

表 1 动态共乘相关研究约束条件分类

Articles	Method	Author	Time Window	Detour Cost
SIGMOD'13 ^[31] ; VLDB'14 ^[32,37]	Noah	Huang Yan, et al.	✓	✓
EUR J OPER RES'14 ^[39]	People and parcels sharing taxis.	Li Baoxiang, et al.	✓	✓
ICTAI'14 ^[38]	Minimising the Driving Distanc	Armant V and Brown K N	✓	×
WISE'14 ^[33]	MARS	Shemshadi A, et al.	✓	×
ICDE'12 ^[4]	T-Share	Ma Shuo, et al.	✓	×
ITSC'12 ^[34]	Distributed Taxi-Sharing System	d'Orey P M, et al.	✓	×
TRANSPORT RES B-METH'11 ^[28]	A simulation study in metro Atlanta	Agatz N A H, et al.	✓	×
IJCAI'11 ^[29]	Parallel Auctions	Kleiner A, et al.	×	×
MATES'09 ^[40]	SMIZE	Xing Xin, et al.	×	✓
EDBT'08 ^[30]	Highly scalable trip grouping	Gidofalvi G, et al.	×	×
EUR J OPER RES'06 ^[35]	A two-phase insertion technique	Coslovich L, et al.	✓	×
PARALLEL COMPUT'04 ^[36]	Parallel Tabu search heuristics	Attanasio A, et al.	✓	×

Notes: “✓” means concerned; “×” means unconcerned.

2.4 动态共乘优化目标

传统静态共乘中,车辆出发前完成车辆与乘客的匹配,车辆出发后行程不再变化,对匹配算法耗时敏感性不高;而动态共乘中,由于新增的行程请求会对已有的行程安排产生影响,进而可能影响到车上已有乘客的行程体验,在进行新的请求匹配时,不仅要考虑到新的共乘请求是否满足约束,还需考虑新增乘客对已接受请求乘客的影响. 动态共乘的优化目标主要有:1)降低车辆总行程开销;2)提高共乘匹配成功率. 在不同的应用场景和研究中优化目标各不相同.

2.4.1 降低车辆行程开销

对某一车辆的路程开销 $W(Car) = \sum_i W(v_i, v_{i+1})$, 优化目标为 $\min W(Car)$. 实时动态共乘的车辆在收到新的请求时,一部分请求位置点已经完成,只需要对尚未完成的请求位置进行规划. 对某一个车辆,可以把该车辆已确认的请求分为 2 类:1)乘客已在车上,乘客对应请求为 tr_i ,此时,车辆已驶过其起点 $tr_i.o$, $schedule$ 中只需对其终点 $tr_i.d$ 进行处理;2)对应 tr_i 请求的乘客还未上车, $schedule$ 中需

对起点 $tr_i.o$ 和终点 $tr_i.d$ 都进行处理. 动态共乘要求车辆在接受请求的时刻 $V.t$,对已在车上乘客的终点 $tr_i.d$,与未上车乘客的起点 $tr_i.o$ 和终点 $tr_i.d$ 所组成的集合 $schedule = \langle v_0, v_1, \dots, v_k \rangle$ 进行重新规划,使得新请求和已接受请求在满足时间和空间约束的情况下,本车的总体行程开销最小. 如图 4 所示,车辆 Car_i 已确定接收 tr_1, tr_2, tr_3 的用户请求,在行驶至点 P 时收到新的用户请求 tr_4 ,此时对应 tr_1 乘客已经在车上,但车辆尚未到达 $tr_1.d$,对应 tr_2 的乘客已经完成输送, tr_3 还未上车,此时,对新的行程请求 tr_4 ,判断其是否能被接受,需要在 P 点对 $\langle tr_1.d, tr_3.o, tr_3.d, tr_4.o, tr_4.d \rangle$ 进行路线规划,判断能否得到符合车上已有乘客请求 tr_1, tr_3 和新的请求 tr_4 约束的路线,并得到最低开销行程作为车辆下一步的行驶路线.

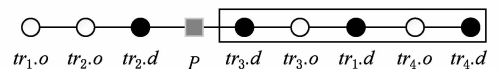


Fig. 4 Rescheduling with a new request^[37]

图 4 收到行程请求需处理的新的规划^[37]

2.4.2 提高共乘匹配成功率

当处理 n 辆车、 m 个服务请求时,还有一个总体目标:匹配成功率.

定义 10. 匹配成功率. 定义 n_{r_suc} 表示在某区域内,在时间窗口 $[time_i, time_j]$ 被有效分配的乘客请求; n_{r_req} 表示在该区域内,在时间窗口 $[time_i, time_j]$ 中所有被提出的请求. 匹配成功率表示在某区域内,某时间窗口中被有效分配的请求数与总请求数的比例,即 $Ratio_{suc} = \frac{n_{r_suc}}{n_{r_req}}$.

动态共乘中乘客请求的匹配效果也是大数据背景下共乘能否可持续发展的重要因素,如乘客的请求不能被有效匹配,会使用户体验下降,而多次匹配失败很可能造成用户流失,匹配成功率可以从宏观尺度上反映匹配效果. 实时动态共乘可近似看成贪婪匹配问题. 乘客发出行程需求后,希望在一定的时

间内得到服务响应. 利用匹配算法,在某个时刻会产生乘客和车辆在这一时刻的最优匹配安排,并将用户请求分配到指定车辆,但随着新的请求加入后,之前“最优安排”的匹配在之后的时刻可能已非最优. 然而对实时的请求处理,这种方法从技术角度容易实现,也容易被乘客和司机所接受,因此对动态共乘的匹配一般都认为是到某一时间点所能达到的最优匹配.

综上,动态共乘主要有总体开销最小化、响应时间最小化、匹配成功率最大化等优化目标:

总体开销最小化,即:

$$\min W(Car), \quad (4)$$

匹配成功率最大化,即:

$$\max Ratio_{suc}, \quad (5)$$

近年来,动态共乘的研究工作的优化目标也不尽相同,总结起来如表 2 所示.

Table2 Optimization Goal of Dynamic Ride Sharing

表 2 动态共乘相关研究优化目标分类

Articles	Method	Author	Total Cost	Matching Rate
SIGMOD'13 ^[31] ; VLDB'14 ^[32,37]	Noah	Huang Yan, et al.	√	×
EUR J OPER RES'14 ^[39]	People and parcels sharing taxis.	Li Baoxiang, et al.	√	×
ICTAI'14 ^[38]	Minimising the Driving Distanc	Armant V and Brown K N	√	×
WISE'14 ^[33]	MARS	Shemshadi A, et al.	√	×
ICDE'12 ^[4]	T-Share	Ma Shuo, et al.	√	×
ITSC'12 ^[34]	Distributed Taxi-Sharing System	d'Orey P M, et al.	√	×
TRANSPORT RES B-METH'11 ^[28]	A simulation study in metro Atlanta	Agatz N A H, et al.	√	√
IJCAI'11 ^[29]	Parallel Auctions	Kleiner A, et al.	√	√
MATES'09 ^[40]	SMIZE	Xing Xin, et al.	×	×
EDBT'08 ^[30]	Highly scalable trip grouping	Gidofalvi G, et al.	×	√
EUR J OPER RES'06 ^[35]	A two-phase insertion technique	Coslovich L, et al.	√	×
PARALLEL COMPUT'04 ^[36]	Parallel Tabu search heuristics	Attanasio A, et al.	√	√

Notes: “√” means concerned; “×” means unconcerned.

3 动态共乘问题分类

动态共乘问题按共乘的完成方式与车上乘客的数量可分为单车辆单乘客、单车辆多乘客、多车辆单乘客、多车辆多乘客 4 种^[41]. 不同类型的共乘模型适用于不同的应用场景,采用的优化方法也不同,本节对动态共乘的不同分类进行介绍. 其中,单车辆多乘客模型最为典型,是本文的重点.

3.1 单车辆单乘客动态共乘

单车辆单乘客动态共乘模型是动态共乘模型中

的基础模型,在这种模型中每台车辆只允许与一名乘客进行共乘匹配,匹配过程中车辆 $Car = \langle id, t, loc, schedule, TR_{rec}, n_{cap}, n_{emp} \rangle$, TR_{rec} 仅有车辆已完成乘客匹配和无乘客 2 种状态. 在这种模型中,默认将 $n_{cap} = 1$,当 $n_{emp} = 0$ 时,车辆不再接受新的请求. 不需要考虑新发出请求的乘客与已经在车上乘客原请求的冲突,只需要考虑乘客与司机的匹配. 在单乘客模型中,司机会绕路去搭载乘客,因此会考虑绕路所带来的开销,Agatz 等人^[28]研究的优化方法中,以最小化总体行程及每个人的花费为目标,约束条件为总的绕路距离小于司机与乘客各自未进行共乘开销

的总和. 作者基于 2008 年 Atlanta 的地铁出行数据进行了动态共乘的出行模拟, 并通过优化方法取代了简单贪心算法, 改进了共乘匹配效果. 而 Kleiner 等人^[29]的研究中, 通过计算司机绕路的最小成本, 在满足最小成本的基础上, 乘客通过价格竞拍的方式与车辆进行匹配. 单车单乘客动态共乘模型是动态共乘算法的基本模型, 比较容易实施, 匹配复杂度低; 但因为一般车辆都有多个座位 ($n_{cap} > 1$) 可提供, 这种共乘模型没有充分发掘运输潜力, 所以很多研究都在单车单乘客模型的基础上进行了扩展.

3.2 单车多乘客动态共乘

在单乘客动态共乘模型的基础上, 单车多乘客动态共乘模型中车辆 $Car = \langle id, t, loc, schedule, TR_{rec}, n_{cap}, n_{emp} \rangle$, 只要 Car 满足车辆上仍有空闲座位, 即 $n_{emp} < n_{cap}$ 都可以接受新的请求, 而对新的请求, 除了满足新的乘客等待时间、绕路距离等服务请求, 还要考虑接新的乘客上车所产生的绕路及上车之后新的 $schedule$ 对原来车上已有乘客行程 TR_{rec} 产生的影响. 如图 5 所示, 插入新的乘客之前, 车辆到达 $v_0, v_1, v_2, v_3, \dots, v_n$ 的时间分别为 $t_0, t_1, t_2, t_3, \dots, t_n$, 在车辆到达 v_2 点之前接受了新的行程 v' 到 v'' , 进行重新规划, 原来到达 $v_0, v_1, v_2, v_3, \dots, v_n$ 的时间序列变化为 $t_0, t_1, \hat{t}_2, \hat{t}_3, \dots, \hat{t}_n$, 且 $\hat{t}_2 > t_2, \hat{t}_3 > t_3, \dots, \hat{t}_n > t_n$, 而每个节点的到达时间在加入新乘客之前是满足乘客约束的, 而如第 2 节所述, 对新的用户请求, 需对已有 $schedule$ 中的位置点和新乘客的位置点进行规划, 判断是否能得到符合所有乘客约束的路线规划, 以确定新的乘客请求能否接受.

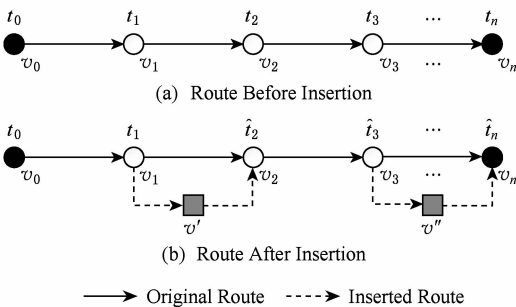


Fig. 5 Insertion in single driver multiple passengers

图 5 单车多乘客共乘插入行程示意图

3.3 多车辆单乘客动态共乘

在实际应用中, 松弛完成某一乘客行程请求的车辆数量, 由原来的一台车辆完成某一乘客的输送变为可由多台车辆协作完成对某乘客的输送, 得到单乘客多车辆模型^[40]. 如图 6 所示, 某时刻 Car_1 , Car_2 已分别与 tr_1, tr_2 匹配, 行程请求 tr_3 出现, 在

单车模型中, 无论 tr_3 与 Car_1 还是 Car_2 匹配, 都会产生较大绕路开销. 而在多车辆单乘客模型中, 首先由 Car_1 将 tr_3 带到指定的换乘点 P_t , 由 Car_2 完成余下的行程, 则可以较小绕路开销完成匹配. 这种模型更为灵活, 提高了用户的匹配成功率, 使资源利用潜力进一步被发掘.

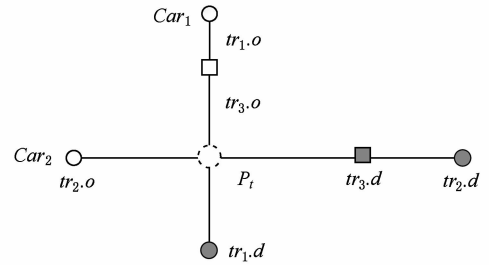


Fig. 6 A composite route offered as a transport solution^[40]

图 6 多车辆单乘客组合路线示意图^[40]

3.4 多车辆多乘客动态共乘

多车辆多乘客模型可以看成是多车辆单乘客动态共乘模型的进一步扩展, 即利用多个车辆进行多个乘客的输送. 对这种模型扩展, 还可以松弛车辆行驶中的司机身份限制, 即在行程中可切换司机, 在某段路上原来充当乘客的人可以充当驾驶员驾驶车辆. 这种模型可以应用于租赁汽车中, 如可随时租赁使用的 ZipCar^[42], 车辆的租用者可以通过智能规划、角色互换实现出行成本的最优化.

4 动态共乘优化主要方法

如 2.3 节和 2.4 节所述, 动态共乘匹配算法需要在满足时间窗口、绕路等约束条件下对总体开销、响应时间、匹配成功率进行优化. 对动态共乘而言, 需要解决共乘匹配时间开销大与动态共乘应用的实时性需求之间的矛盾, 解决这种矛盾的经典方法是利用 Filter and Refine 框架. 本节首先介绍了大数据背景下动态共乘优化问题, 随后围绕 Filter and Refine 框架对动态共乘优化方法进行详细阐述.

4.1 动态共乘优化问题的复杂性

动态共乘的实时处理是其大数据背景下广泛应用的挑战之一. 以司机的角度, 新的请求到来时, 司机可能已经接受了一些乘客行程请求, 并在接受新的行程请求时已对这些行程请求安排了行驶路线. 对一个新的请求, 车辆需要重新进行路线安排以判断其是否能接受新的请求. 如 2.4.1 节所述情景, 车辆行驶至点 P 时, 已经接受 tr_1, tr_2, tr_3 用户请求, 并

完成乘客 tr_2 的输送,如图 4 所示,对应 tr_2 乘客已经在车上,已经到达 tr_1, o 但尚未到达 tr_1, d , 对应 tr_2 的乘客已经完成输送,而 tr_3 还未上车,需要行驶经过 tr_3, o, tr_3, d , 对应于 tr_4 , 需行驶至 tr_4, o, tr_4, d . 如图 7 所示,确定车辆能否接受新行程请求 tr_4 需在 P 对节点 $\langle tr_1, d, tr_3, o, tr_3, d, tr_4, o, tr_4, d \rangle$ 进行路线规划,判断能否找到同时满足:①车上空位约束 $Car.n_{emp}$;②时间窗口 $\langle tr_1, t_{ow}, tr_3, t_{ow}, tr_3, t_{dw}, tr_4, t_{ow}, tr_4, t_{dw} \rangle$ 约束;③每名乘客的行程绕路约束 $\delta_{det}(tr_1, o, tr_1, d), \delta_{det}(tr_3, o, tr_3, d), \delta_{det}(tr_4, o, tr_4, d)$ 路线规划 $schedule = \langle v_0, v_1, \dots, v_k \rangle$. 该规划过程可看成图搜索问题,是拓展的旅行商问题,已被证明为 NPC^[4,37] 问题.

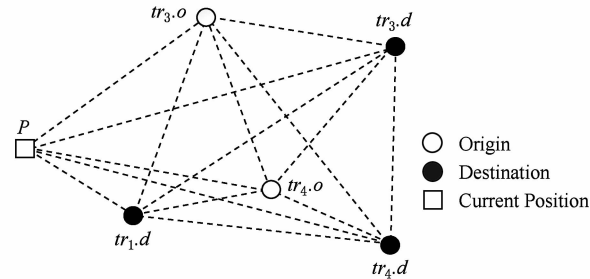


Fig. 7 Possible schedules with a new request

图 7 行程规划示意图

4.2 动态共乘优化问题实时性需求

定义 11. 响应时间. $Time_{comp} = tr_i, t_{get} - tr_i, t_{req}$, 其中, tr_i, t_{req} 表示服务请求发起时间, tr_i, t_{get} 表示系统判断新请求能否被成功匹配的时间. 降低响应时间以满足动态共乘的实时性需求, 是决定大数据背景下动态共乘应用效果的关键技术之一.

如 4.1 节所述, 车辆接受到新的请求后, 需在 P 点对 $\langle tr_1, d, tr_3, o, tr_3, d, tr_4, o, tr_4, d \rangle$ 进行规划, 以期在满足约束条件下得到最优化的 $schedule = \langle v_0, v_1, \dots, v_k \rangle$. 动态共乘的实际应用中, 车辆在路网上移动, 如果不能在有效的时间内对用户请求进行相应匹配, 则无法满足动态共乘匹配的实时性要求. 因此需设计专门的优化算法来满足动态共乘的实时性需求. 减少用户请求响应时间是动态共乘大规模应用需解决的关键技术, 后文将主要围绕动态共乘实时性的有关研究进行展开.

4.3 动态共乘优化的 Filter and Refine 框架

近年来对大数据背景下动态共乘的研究中, 解决实时性问题主要采取 Filter and Refine 框架^[43]. 基本思想是: 首先利用某些过滤条件将不能完成匹配的车辆或乘客从车辆集合 C 和乘客请求集合 TR

中剔除, 仅保留可能符合条件的车辆或乘客请求, 或通过聚合分组等方法缩小问题解空间; 然后在较小的解空间中对车辆和乘客进行匹配. 即:

Step1. Filter. 给定 $G = \langle V, E, W \rangle$ 、服务请求集合 $TR = \langle tr \rangle$ 、车辆集合 $C = \langle Car \rangle$, 对 $tr_i \in TR$, 将 C 中无法满足匹配约束的车辆进行过滤, 得到 $C_{Filtered} = \langle Car \rangle$, $|C_{Filtered}| \leq |C|$; 对 $Car_i \in C$, 将 TR 中无法满足的行程请求进行过滤, 得到 $TR_{Filtered} = \langle tr \rangle$, $|TR_{Filtered}| \leq |TR|$; 通过聚合等方法将 TR 聚合成小数量的请求组, 得到 $TR_{Grouped} = \langle tr \rangle$, $|TR_{Grouped}| \leq |TR|$.

Step2. Refine. 将 Filter 步骤获得的候选车辆集合 $C_{Filtered}$ 中的车辆, 与乘客请求集合 $TR_{Filtered}$ 进行匹配, tr 满足上下车时间窗口 t_{ow}, t_{dw} 和绕路开销 δ_{det} 等约束的基础上, 得到行程开销最小的乘客-车辆匹配对.

4.4 Filter 主要方法

Filter 步骤中, 目标是通过剪枝与过滤降低 Refine 中的候选解规模. 动态共乘, 乘客发出请求后, 需要对在路网上行驶的车辆进行匹配检索, 针对移动目标快速索引问题, 虽然已有一些动态空间索引技术^[44-45], 但这些算法不是专门为动态共乘设计的. 如 R-tree 结构, 因为多目标的位置不断变换, 会引起叶子节点的不断更新, 带来额外开销. Filter 步骤主要有预先计算可行解、建立动态空间索引、利用价格开销过滤和基于分组预处理聚类等方法.

4.4.1 预先计算可行解的优化方法

如前所述, 在电话叫车服务中的静态共乘服务, 匹配在车辆出发前进行, 对实时性要求不高, 基于离线操作可得到较好的优化效果, 然而对于动态共乘, 因不断有新请求到来, 需进行实时规划, 静态匹配的方法已经不能满足实时性需求. Coslovich 等人^[35] 提出 2 阶段插入算法改进系统的实时性: 第 1 阶段, 利用乘客请求的时间间隙进行预先计算, 得出当前路线中车辆可有效响应的路线安排, 在这一步计算中, 对 n 个可停靠的站点, 根据搜索可得到的有效路线数量由少变多设计了 2 种算法, 复杂度分别为 $O(n^3)$ 与 $O(n^4)$. 实际应用时, 首先采用搜索到路线少、复杂度为 $O(n^3)$ 的可选路线搜索方法, 在时间充裕的情况下调用搜索到可选路线多、复杂度为 $O(n^4)$ 的方法. 第 2 阶段, 当动态的请求发生时, 通过第 1 阶段的预先计算的可行解, 对新的行程请求进行判断, 并引进局部松弛时间和全局松弛时间对新的请求进行加速, 在这个阶段插入的时间复杂度为 $O(n)$. 这种优化方法面向传统的电话叫车服务, 适用于中小

规模的动态共乘问题.

4.4.2 建立动态空间索引优化方法

车辆在路网上行驶,快速检索出可能符合约束条件的候选车辆,缩小规划与匹配的候选车辆数,可有效降低共乘系统响应时间.郑宇等人^[4]针对出租车共乘提出了基于区块的动态空间索引方法.如图8利用空间索引方法优化匹配示意图^[4]所示,这种方法预先将空间分成小的区块,并根据区块内部的路网分布,如图8(a)所示,将单元区块用路网上靠近区块中心的点来表示,计算基于路网区块距离,

生成区块距离矩阵,如图8(b)所示,每个区块记录了时间近邻区块索引、空间近邻区块索引、区块内将会到达出租车的索引,其中,时间近邻区块索引和空间近邻区块索引按时间与空间的递增排序,出租车索引按到达时间排序,如图8(c)所示.用户发出乘车请求 $tr = \langle o, d, p_n, t_{ow}, t_{dw}, \epsilon \rangle$ 后,分别从请求的起点 $tr.o$ 和终点 $tr.d$ 所在方格 $grid_i$ 和 $grid_j$ 按照距离由近到远的方式在空间近邻索引中拓展搜索区块,得到近邻区块中的出租车作为候选车辆,完成候选车辆的第1步过滤,如图8(d)所示.

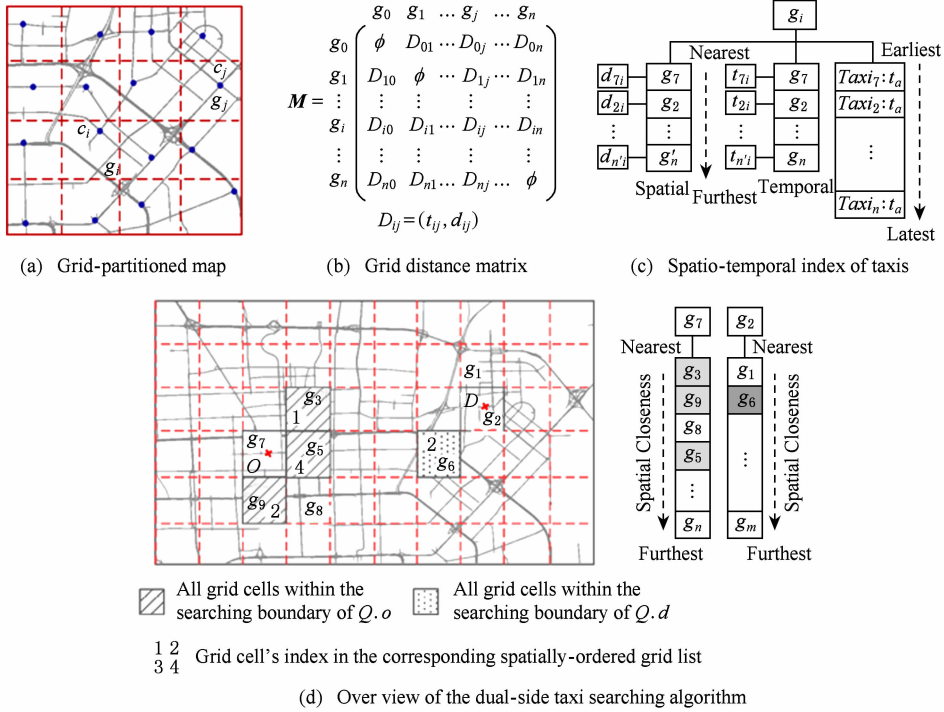


Fig. 8 A spatio-temporal index to quickly retrieve candidate taxi^[4]

图8 利用空间索引方法优化匹配示意图^[4]

4.4.3 基于请求分组预处理及并行优化方法

大数据背景下的大规模共乘问题中,动态的行程可以看成是流数据的分组问题^[46],即把具有相近距离起点、终点的请求进行分组,降低 Refine 步骤的解空间.Gidofalvi 等人^[30]针对大规模行程请求,实现了基于 SDSQ 流数据管理系统的行程分组 (trip grouping, TG) 并行化算法,在 TG 算法中用额外开销 FEC (fractional extra cost) 来衡量不同请求行程间的关系,给定 2 个请求 tr_i 与 tr_j :

$$FEC(tr_i, tr_j) = \frac{\text{dist}(tr_i.o, tr_j.o) + \text{dist}(tr_i.d, tr_j.d)}{\text{dist}(tr_i.o, tr_i.d)}$$

某一请求 tr_i 对已经聚合请求集合 s 的平摊开销定义为 AC (amortized cost):

$$AC(r_i, s) = \frac{1 + \sum_{r_j \in s} FEC(tr_i, tr_j)}{|s|}$$

通过计算行程之间的 FEC 以及行程请求和行程请求组的 AC, TG 算法实现了对请求进行分组.如图9所示,对请求 tr_1, tr_2, tr_3 :

$$FEC(tr_1, tr_2) = \frac{\text{dist}(tr_2.o, tr_1.o) + \text{dist}(tr_2.d, tr_1.d)}{\text{dist}(tr_1.o, tr_1.d)}$$

$$FEC(tr_1, tr_3) = \frac{\text{dist}(tr_3.o, tr_1.o) + \text{dist}(tr_3.d, tr_1.d)}{\text{dist}(tr_1.o, tr_1.d)}$$

$$AC(\{tr_1, tr_2, tr_3\}) = \frac{1 + FEC(tr_1, tr_2) + FEC(tr_1, tr_3)}{3}$$

m 个请求未做并行化处理的算法复杂度为 $O(m^3)$. 实现并行计算时,不同的任务划分方法对并行计算效果影响较大,该文分别测试了基于静态点的 4 分法 SPQ(static point quad partitioning)、基于中位数的静态多维划分方法(static KD partitioning, SKD)、自适应的基于点的 4 分法(adaptive point quad partitioning, APQ)、自适应多维划分方法(adaptive KD partitioning, AKD) 四种数据流的划分方法,通过实验得出自适应多维划分方法具有最好的划分效果,在 16 个核的机器上进行并行计算时,较串行算法获得了 10 倍以上的加速比,为请求分组及利用并行计算的方法进行共乘数据处理都提供了一定借鉴.

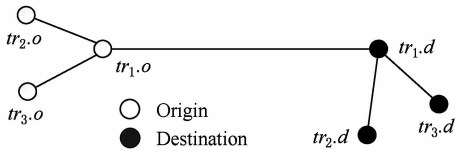


Fig. 9 Illustration of fractional extra cost(FEC) and amortized cost(AC) calculations w. r. t request tr_2 [30]

图 9 额外开销(FEC)与平摊开销(AC)计算示意图[30]

基于分布式系统对移动物体进行检索的研究近年也取得了相关进展,于自强等人[47]设计了动态条带索引(dynamic strip index, DSI)的索引结构,与基于方格的区块划分不同,动态条带索引采用条带划分的方法对移动物体进行索引,保证了数据的分布式处理效果,同时作者基于 DSI 设计了分布式移动物体的 K 近邻检索算法(distributed k -NN search, DKNN),并在 Apache S4 上进行了实现,获得了较好的可扩展性与查询速度,为利用分布式系统的方法来提高移动物体索引效率提供了一定借鉴.

4.5 Refine 主要方法

如 4.4 节所述,在经过 Filter 步骤处理后,需匹配的车辆集合 C 与行程请求集合 TR 空间已经缩小. Refine 步骤,需要对 Filter 步骤中获得的候选车辆集合和新的请求集合进行匹配. 能否将新的请求与车辆进行匹配,取决于新请求和车辆上已有的 TR_{rec} 进行路线规划后能否找到满足约束的可行解. 如 4.1 节所述对车辆路线的规划是一个 NPC 问题,如果利用精确算法,解决匹配问题的时间会随着问题规模的增大而呈现指数级的增长. 在路径规划的研究中,比较典型的有分支定界法[48]、整形规划法[49]、遗传算法和蚁群算法等方法,但这些方法主要解决的问题是离线匹配,远不能满足实时动态需

求. 因此针对动态共乘,需专门地优化算法来加快行程规划速度. 现有动态共乘 Refine 步骤主要优化手段有简化计算模型、引进专有数据结构、采用启发式方法等.

4.5.1 简化计算模型的优化方法

为满足实时性, T-share[44] 在利用调度算法检测候选出租车是否符合要求时,采取了简化问题模型方法,即对新的请求不重新规划路线,仅在原行程安排中插入新请求,将规划问题精简为插入判断问题,同时引入了距离矩阵和松弛时间,利用距离三角不等式来对新插入行程的合法性进行判断,对有 n 个点的行程,将算法复杂度由 $O(n!)$ 降为 $O(n^2)$. 这种方法提高了动态共乘系统的实时性,但因为简化了问题,会对解的最优性造成损失.

4.5.2 利用新型树结构的优化方法

黄艳等人[37]提出了基于活动树(Kinetic Tree)数据结构的优化方法. 一般的分支定界法对新请求的判断未利用原来的计算结果,从而产生了重复计算,导致运算时间增加. 作者为了避免了重复运算,设计了活动树. 如图 10 所示,活动树中保存了当前所有可行规划,根节点到叶子节点的次序表示规划路径. 新的乘客请求到来时,如果采用图搜索的方法,每一步都需对多点图进行搜索,无法利用之前的计算结果,算法复杂度与节点处呈几何级数增长. 而通过活动树保存当前最优的路径安排和所有有效安排 $schedule = \langle v_0, v_1, \dots, v_k \rangle$, 在收到新的请求时,只需基于保存在活动树中的有效路线对新请求 $tr.o, tr.d$ 进行插入有效判断即可. 如图 10 所示,令 s_i 表示行程 tr_i 请求的起点 $tr_i.o$, 令 e_i 表示行程 tr_i 请求的终点 $tr_i.d$. 收到第 1 个请求生成如图 10(a) 所示活动树. 当第 2 个乘客请求到达时,第 1 个乘客已经上车. 此时利用 10(a) 所示活动树对乘客 2 的起、终点有效性进行判断,假设乘客 1 先下车再接乘客 2 路线不满足约束,得到先送乘客 2 到终点和先送乘客 1 到终点这 2 个可能路线,其中最优路线为 (l, s_2, e_1, e_2) , 而符合约束的可行路线 (l, s_2, e_2, e_1) 也被记录,生成如图 10(b) 所示活动树. 在车辆向乘客 2 起点行驶途中,收到乘客 3 请求,此时需要判断先接乘客 2 还是乘客 3,假设此时有 5 条可满足约束的路线,其中最优路线为 $(l, s_2, s_3, e_2, e_3, e_1)$, 得到如图 10(c) 所示活动树. 乘客 2 上车后,收到乘客 4 请求,此时图 10(c) 中 r_3 右子树已无效,假设此时有 2 种路线满足共乘约束,则得到如图 10(d) 所示活动树. 同时作者针对路线中节点过多造成空间搜索

开销过大的问题,提出了基于热点的优化方法,将距离小于阈值 θ 的点聚合为一个 hotspot 节点,使得插入位置判断次数减少,进一步缩小解空间,降低系统运算量,提高响应速度.作者利用上海实际出行数据进行了实验仿真与性能测试,实验表明与整形规划方法相比,活动树方法获得了 1 个数量级以上的加速比.

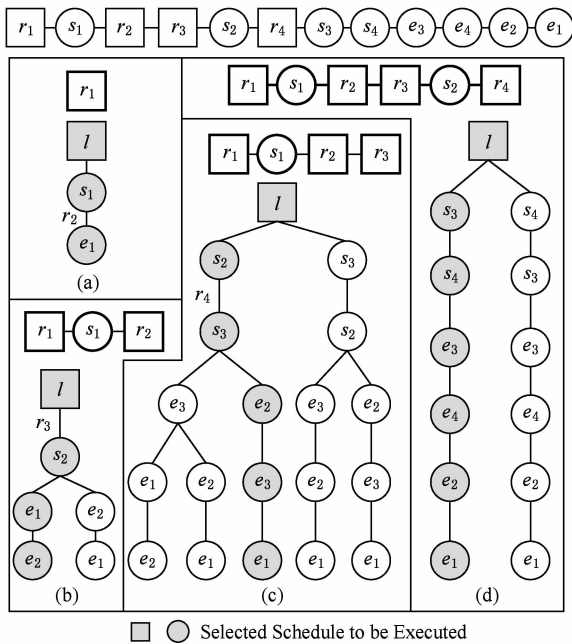


Fig. 10 Kinetic Tree for trip schedules^[37].

图 10 利用 Kinetic Tree 方法优化行程规划示意图^[37]

4.5.3 利用启发式算法的优化方法

对应于路线的规划问题,遗传算法、蚁群算法、禁忌搜索算法、模拟退火方法、粒子群优化等启发式优化方法也被采用. Herbawi 等人^[50]比较了蚁群算法和遗传算法在动态共乘中的性能,并基于带有时间窗口的多乘客多车辆动态共乘问题,提出了遗传算法的解决方案^[51]. 在多乘客多车辆动态共乘问题中,优化目标为车辆行程时间最短、车辆行驶距离最短、乘客等待时间最短以及最大化的匹配效果. 利用遗传算法进行行程安排时,每一个车辆的行程被记录到列表中 $schedule = \langle v_0, v_1, \dots, v_k \rangle$, 同时未被满足的乘客请求也被记录到一个不满足列表中. 遗传算法初始化时,首先将车辆的起终点插入到车辆的行程安排中,而后随机选择路线安排插入已有路线中,作为遗传的父代. 在生成可行解后,选取可行路线并对其中的路线通过交换来生成后代,作为单点交叉算子,该文对路线中的路径点提出了 5 种变异操作. Santos 等人^[52]提出了面向出租车的动态共乘

优化问题框架和启发式算法,在该框架中以乘客和车主的总成本最小为优化目标,以插入行程给原车辆带来额外的时间开销作为贪心函数,并采取了本地搜索优化的方法优化匹配效率,该文的作者对数百个真实请求数据进行了动态共乘匹配模拟,得到了秒级的响应时间.

5 动态共乘人文影响

动态共乘将运力资源进行了整合优化,这种资源整合过程,除了匹配与路线规划计算本身,司机与乘客的社会行为因素也是共乘系统需考虑的因素. 本节对影响乘客体验的信用体系、价格机制和用户信息交互等人文因素在共乘方面所进行的研究进行阐述.

5.1 信用体系

共乘形成于熟人之间,而后随着通信与交互手段逐步发展^[11-12],这种行为本身离不开人们之间的信任关系. Chaube 等人^[53]对大学校园共乘系统研究表明,共乘参与者之间的信任关系直接影响共乘参与度,陌生人之间有 7% 的人接受共乘,在熟人之间有 98% 接受共乘,而与熟人的朋友的共乘也会有 69% 被接受. 因此, Cici 等人^[54]通过分析用户电话呼叫记录,发掘共乘潜力. Facebook, Twitter、微博、微信等社交应用把兴趣爱好、背景相同的人形成了一种天然的聚合. 利用社交网络信息进行用户爱好分析,增加用户的参与度已经被应用在实现系统中. ZimRide, Avego 和 Carticipate 等系统已经开始利用 SNS 进行用户的连接,而 Uber 也在推出熟人共乘补贴等服务,培养用户共乘习惯. 为增加用户认同感, Blablacar 等共乘系统甚至把司机的聊天水平作为评价体系的一部分. 但大数据背景下的动态共乘中,司机与乘客的关系并不仅限于熟人,因此在共乘系统中,需要一个完善的信用体系,规避可能出现的信用风险,提高司机服务质量和乘客的参与热情,使共乘系统长期、稳定、健康发展.

通过信用体系来保证用户长期参与性也是很多 C2C 和 O2O 系统中需要解决的问题^[55-56]. 我们认为,共乘系统中信用体系需考虑信用可查、信用可评、信用公平、失信处罚等因素^[56]. 在建立信任的过程中,历史交易积累产生的信任感是最直接有效的. 很多共乘系统采取了第三方支付+信用评价的方式,将用户反馈与第三方支付系统相结合促进交易经验分享,保证交易资金安全,约束乘客和司机的行

为,产生良性循环,使共乘系统更好地为乘客和司服务。

第三方支付^[57]中钱款不直接在司机与乘客间流动,采用第三方信用担保的方式,保证支付安全,如Paypal、支付宝等。第三方支付的典型流程如下:在完成乘客和司机匹配后,乘客将预付费放到第三方支付托管机构;在乘客按预先约定好的方案完成行程后,对司机服务进行评价,在确定完成服务后,第三方平台将费用转给司机,从而完成整个共乘的支付。为保证支付安全和信用体系的构建,对第三方支付平台不仅要求其进行资金托管,更要求其具有权威、独立的评价能力,能对用户产生足够的约束力。当前Avego、Zebigo、Goloco、51用车、嘀嗒拼车、天天拼车等系统采用的都是这种支付方式。

用户反馈机制的构建,往往采用不同维度的互评体系,共乘完成后由乘客和司机互相评价。我们认为动态共乘信用评价体系需满足3项需求:①保证司机与乘客历史信用的可查询性,司机和乘客可以在匹配前对共乘人员进行了解;②建立完善的失信处罚机制,如某个乘客在确定匹配后多次单方面退出行程,那么在下次匹配的过程中,将对其匹配的优先级做一定的惩罚,甚至取消其共乘匹配资格;③保证信用评价的公平性,充分考虑影响服务质量的因素的多元性。如某车辆已载有1名乘客,司机与第2名乘客达成了一个新的共乘请求后,第3名的共乘请求到达,司机也接受了第3名乘客的请求,在司机接第2名乘客上车时,由于第2名乘客没有按时到达指定地点,造成了之后对第3名乘客的服务延迟。如果只是简单的评分,必然会影响司机信用等级的公平性。因此针对共乘出行模式本身的特点,制定专有的信用评价体系对共乘系统广泛推广有重要作用。

5.2 价格机制

根据乘客出行费用与司机收入的决定方式,目前共乘价格机制有按行程比例计算、通过拍卖竞价等。Agatz等人^[28]以单乘客的模式为研究对象,乘客费用与其行程距离成正比,郑宇等人^[4]在T-share系统中也采用了类似的价格系统。动态匹配问题中用户的参与度也可通过制定合理的价格机制来保证。Kleiner等人^[29]为保证匹配的接单率,提出了基于拍卖的激励模式,该模式中以第2高的价格作为竞拍成交价,与其他的动态共乘系统在减少总的行程上的目标不同,该文在满足不同参与者的个人需求和最小化车辆行驶距离的基础上,以提高匹配成功率为目标,提升用户体验,保证系统长期可用性。

如图11所示 $lp(p) = dist(tr_2, o, tr_2, d)$ 表示乘客共乘前的路网最短距离; $ld(d) = dist(tr_1, o, tr_1, d)$,表示司机原行程中最小行驶距离; $lt(d, p) = dist(tr_1, o, tr_2, o) + dist(tr_2, o, tr_2, d) + dist(tr_2, d, tr_1, d)$,表示乘客在共乘之后的最小行驶距离。司机每公里的成本为 C_{km}^d ,结合司机绕路的总行程和原来的行程,得到司机行程的总估值为: $V_d(\langle d, p \rangle_i) = (ld(d) - lt(d, p)) \times C_{km}^d$,作为司机本身绕路后的成本价格,而乘客给出的拍卖价格如果大于这个价格,即 $V_p(\langle d, p \rangle_i) \geq ld(p) \times C_{km}^d$ 时,则共乘的匹配是有效的。

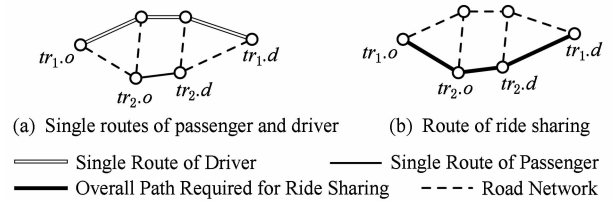


Fig. 11 Total cost for a ride sharing with detour^[29]

图 11 共乘成本开销示意图^[29]

Kamar等人^[58]提出了一个基于VCG(Vickrey-Clarke-Groves)^[59]机制的定价策略,利用该机制保证定价的激励相容等特性,同时该文作者测试了不同策略对计算效率、预算平衡、激励相容、匹配成功率的影响。Cheng等人^[60]提出了解决生活中最后一公里到达的共乘方案,考虑了动态和需求响应机制来安排非专用商业车队(出租车或多乘客小客车)共乘,提出了基于征求意见支付率(willing payment rate)和意愿补偿率(compensation rate)的竞拍机制,并基于现实和模拟数据测试了在2种竞拍机制下不同参数对共乘的影响。Zhao等人^[61]提出了一个基于市场调节的共乘系统,在这个系统中通勤者通过他们提出的出行约束来进行匹配,比如可用的座位数、开销,为增加匹配度,通勤者可充当司机或乘客的角色。在之前的策略中,VCG策略可能会导致系统需要补贴才能正常运转,产生赤字。固定价格的方法不能满足激励相容,作者提出了一个带有双边底价的VCG机制,具有激励相容的特性,并使底价在一定范围内赤字得以控制。

5.3 信息交互与隐私问题

移动设备及可穿戴设备的普及,为有效解决系统与人的快速交互提供新的途径,有效的人机接口对共乘体验起着重要作用。面向共乘的人机交互主要解决以下4个问题^[62]:1)有效信息获取;2)隐私保护;3)易用性;4)智能化信息处理。

随着智能手机的应用,共乘平台由基于 PC 的桌面网页浏览向智能终端迁移.用户通过移动端提交行程的起点、终点及相关约束,较以往更为快捷.共乘系统因优化目标不同,需要用户提交信息也有所差别:如 Noah 系统^[32]是以关心用户本人的乘客感受为目标,乘客在进行共乘匹配时,提交当前地点、目的地、等待时间和绕路约束;T-Share 系统^[4]中用户提交上车地点、下车地点及上车和下车的时间窗口.这 2 种系统都有效获取了用户需求,及时将位置及行程信息提交给系统与车辆进行匹配.在进行匹配定位时,为直观地显示司机与乘客的位置,共乘平台通过调用现有的 Google Map、Bing Map、百度地图、高德地图等第三方地图服务商的接口,实现更直观有效的交互.随着语音识别和文本分析技术的不断发展,交互手段越来越人性化^[63],语音、人脸识别等交互手段被广泛采用.在实际应用中,语义上下文反映了用户的当前应用场景,基于语义场景为用户提供更智能服务的研究^[64]也逐步展开.

随着人机交互、社交网络、城市计算和上下文感知技术的发展,场景感知、用户肖像技术在带来精准智能服务的同时也带来了个人信息安全问题.如地址信息会直接表现出生活空间、兴趣地点等个人信息,而乘客在共乘时也会考虑个人隐私泄露的问题,在信息有效交互的基础上保证乘客和司机的个人隐私也是共乘需要面对和解决的问题,这方面的研究也逐步展开,主流媒体和学术界对个人隐私和安全的关注越发凸显.近年来保护用户隐私的信息交互方法被提出^[65],如共乘系统在匹配成功之前不会泄露用户和司机的地址位置信息,以其他形式进行有效的可视化显示,只有在匹配成功后才告知相关位置,减少用户的隐私泄露程度.

6 共乘问题展望

现在对动态共乘的研究大都还没有考虑到当前实际交通情况对共乘的影响,基于社交网络挖掘共乘潜力的研究也处在初级阶段.我们相信动态共乘技术将向 4 个方面发展:1)通过对社交网络进行研究,吸引更多的人参与到共乘中来,形成良性循环;2)利用历史交通数据和当前交通情况等多元信息的路况预测技术;3)优化复杂路况情况下智能匹配技术;4)依靠环境感知等技术,为乘客提供更为智能易用的共乘服务.

共乘随着参与者数量的增加会产生良性循环,

当参与人数达到一个保证服务的“临界点”后,参与司机和乘客的数量将会出现稳定的自增长.当前,共乘还主要局限于长途和上下班通勤,形式以静态共乘为主,随着参与用户数量的增加与动态共乘技术的发展,动态共乘将会在未来得到广泛应用.而共乘技术也同样适用于物流中的请求处理与规划.有关研究表明,一台自动驾驶汽车可以代替当下 10 辆私家车^[66],无人驾驶技术已日趋完善,其与动态共乘匹配技术结合,将会更有效地解决城市中的交通拥堵问题.共乘技术带来的不仅是交通状况的改进、自然环境的改善、出行成本的降低,更是全新的智能生活方式.

参 考 文 献

- [1] Zhang Xiaoye, Sun Junying, Wang Yaqiang, et al. Factors contributing to haze and fog in China [J]. Chinese Science Bulletin, 2013, 58(13): 1178-1187 (in Chinese)
(张小曳, 孙俊英, 王亚强, 等. 我国雾-霾成因及其治理的思考[J]. 科学通报, 2013, 58(13): 1178-1187)
- [2] Firnkorn J, Muller M. What will be the environmental effects of new free-floating car-sharing systems? The case of car2go in ulm [J]. Ecological Economics, 2011, 70(8): 1519-1528
- [3] d'Orey P M, Fernandes R, Ferreira M. Reducing the environmental impact of taxi operation; The taxi-sharing use case [C] //Proc of the 12th IEEE Int Conf on Its Telecommunications (ITST-2012). Piscataway, NJ: IEEE, 2012: 313-317
- [4] Ma shuo, Zheng Yu, Wolfson O. T-share: A large-scale dynamic taxi ridesharing service [C] //Proc of the 29th IEEE Int Conf on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2013: 410-421
- [5] Dillenburg J F, Wolfson O, Nelson P C. The intelligent travel assistant [C] //Proc of the 5th IEEE Int Conf on Intelligent Transportation Systems(ICITS). Piscataway, NJ: IEEE, 2002: 691-696
- [6] Chan N D, Shaheen S A. Ridesharing in north america; Past, present, and future [J]. Transport Reviews, 2012, 32(1): 93-112
- [7] Badoe D A, Miller E J. Transportation-land-use interaction; Empirical findings in North America, and their implications for modeling [J]. Transportation Research Part D: Transport and Environment, 2000, 5(4): 235-263
- [8] Konishi H, Mun S I. Carpooling and congestion pricing; HOV and HOT lanes [J]. Regional Science and Urban Economics, 2010, 40(4): 173-186
- [9] Ma Shuo, Wolfson O. Analysis and evaluation of the slugging form of ridesharing [C] //Proc of the 21st ACM SIGSPATIAL Int Conf on Advances in Geographic Information Systems. New York: ACM, 2013: 64-73

- [10] Metropolitan Transportation Commission. 511.ORG [OL]. [2015-07-01]. <http://www.511.org>
- [11] Shaheen S, Cohen A. Growth in worldwide carsharing: An international comparison [J]. *Journal of the Transportation Research Board*, 2007, 1992(10): 81-89
- [12] Shaheen S A, Cohen A P. Carsharing and personal vehicle services: Worldwide market developments and emerging trends [J]. *International Journal of Sustainable Transportation*, 2012, 7(1): 5-34
- [13] Yang Jie, Li Xiaoping, Chen Tian. A group mining method for incremental spatio-temporal trajectory big data [J]. *Journal of Computer Research and Development*, 2014, 51 (Suppl2): 76-85 (in Chinese)
(杨杰, 李小平, 陈焱. 基于增量时空轨迹大数据的群体挖掘方法[J]. *计算机研究与发展*, 2014, 51(增刊2): 76-85)
- [14] Zhang Junping, Wang Feiyue, Wang Kunfeng, et al. Data-driven intelligent transportation systems: A survey [J]. *IEEE Trans on Intelligent Transportation Systems*, 2011, 12 (4): 1624-1639
- [15] Carma Technology Corporation. carmacarpool [OL]. [2015-07-01]. <https://carmacarpool.com>
- [16] Comuto. mitfahrgelegenheit [OL]. [2015-07-01]. <http://www.mitfahrgelegenheit.de>
- [17] Zimride c/o Enterprise Holdings, Inc. Zimride [OL]. [2015-07-01]. <https://www.zimride.com>
- [18] Carpooling. Carpooling [OL]. [2015-07-01]. <https://www.carpooling.com>
- [19] BlaBlaCar Technology Corporation. Blablacar [OL]. [2015-07-01]. <https://www.blablacar.com>
- [20] Uber Technologies Inc. Uber [OL]. [2015-07-01]. <https://www.uber.com>
- [21] Lyft, Inc. Lyft [OL]. [2015-07-01]. <https://www.lyft.com>
- [22] Tang Liming, Liu Qihua. Neighborhood carpool practice: A case study on community carpool [J]. *Urban Transport of China*, 2010, 8(6): 29-33 (in Chinese)
(汤黎明, 刘其华. 邻里合乘——社区拼车常态化的探索 [J]. *城市交通*, 2010, 8(6): 29-33)
- [23] Beijing Xiaoju Technology Co, Ltd. 滴滴 [OL]. [2015-07-01]. <http://www.xiaojukeji.com>
- [24] China Auto Rental Holdings Inc. 神州专车 [OL]. [2015-07-01]. <http://zhuanche.zuche.com>
- [25] Beijing Shuailai Century Technology Co, Ltd. 51用车 [OL]. [2015-07-01]. <http://www.51yche.com>
- [26] Beijing Changxing Information Technology Co, Ltd. 嘀嗒 [OL]. [2015-07-01]. <http://www.didapinche.com/home>
- [27] Hall R W, Qureshe A. dynamic ride-sharing: Theory and practice [J]. *Journal of Transportation Engineering*, 1997, 123(4): 308-315
- [28] 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
- [29] Kleiner A, Nebel B, Ziparo V A, et al. A mechanism for dynamic ride sharing based on parallel auctions [C] // *Proc of the 22nd Int Joint Conf on Artificial Intelligence (IJCAI)*. San Francisco: Morgan Kaufmann, 2011: 266-272
- [30] Gidofalvi G, Pedersen T B, Risch T, et al. Highly scalable trip grouping for large-scale collective transportation systems [C] // *Proc of the 11th Int Conf on Extending Database Technology: Advances in Fatabase Technology*. New York: ACM, 2008: 678-689
- [31] Shao Zengzhen, Wang Hongguo, Liu Hong, et al. Heuristic optimization algorithms of multi-carpooling problem based on two-stage clustering [J]. *Journal of Computer Research and Development*, 2013, 50(11): 2325-2335 (in Chinese)
(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, 50(11): 2325-2335)
- [32] Tian C, Huang Yan, Liu Zhi, et al. Noah: A dynamic ridesharing system [C] // *Proc of the 2013 ACM SIGMOD Int Conf on Management of Data*. New York: ACM, 2013: 985-988
- [33] Shemshadi A, Sheng Q Z, Zhang W E. A decremental search approach for large scale dynamic ridesharing [C] // *Proc of Web Information Systems Engineering-WISE 2014*. Berlin: Springer, 2014: 202-217
- [34] d'Orey P M, Fernandes R, Ferreira M, et al. Empirical evaluation of a dynamic and distributed taxi-sharing system [C] // *Proc of the 15th IEEE Int Conf on Intelligent Transportation Systems (ITSC 2012)*. Piscataway, NJ: IEEE, 2012: 140-146
- [35] Coslovich L, Pesenti R, Ukovich W. A two-phase insertion technique of customers for a dynamic dial-a-ride problem [J]. *European Journal of Operational Research*, 2006, 175(3): 1605-1615
- [36] Attanasio A, Cordeau J F, Ghiani G, et al. Parallel tabu search heuristics for the dynamic multi-vehicle dial-a-ride problem [J]. *Parallel Computing*, 2004, 30(3): 377-387
- [37] Huang Yan, Bastani F, Jin Ruoming, et al. Large scale real-time ridesharing with service guarantee on road networks [C] // *Proc of the 40th Int Conf on Very Large Data Bases*. New York: VLDB Endowment, 2014: 2017-2028
- [38] Armant V, Brown K N. Minimizing the driving distance in ride sharing systems [C] // *Proc of the 26th IEEE Int Conf on Tools with Artificial Intelligence (ICTAI 2014)*. Piscataway, NJ: IEEE, 2014: 568-575
- [39] Li Baoxiang, Krushinsky D, Reijers H A, et al. The share-a-ride problem: People and parcels sharing taxis [J]. *European Journal of Operational Research*, 2014, 238(1): 31-40
- [40] Xing Xin, Warden T, Nicolai T, et al. SMIZE: A spontaneous ride-sharing system for individual urban transit [C] // *Proc of the 7th German Conf on Multiagent System Technologies (MATES)*. Berlin: Springer, 2009: 165-176

- [41] Agatz N, Erera A, Savelsbergh M, et al. Optimization for dynamic ride-sharing: A review [J]. *European Journal of Operational Research*, 2012, 223(2): 295-303
- [42] Bardhi F, Eckhardt G M. Access-based consumption: The case of car sharing [J]. *Journal of Consumer Research*, 2012, 39(4): 881-898
- [43] Shen Bilong, Huang Yan, Zhao Ying. Dynamic ridesharing [J]. *SIGSPATIAL Special*, 2015, 7(3): 3-11
- [44] Mokbel M F, Xiong X, Aref W G. Sina: Scalable incremental processing of continuous queries in spatio-temporal databases [C] // *Proc of the 2004 ACM SIGMOD Int Conf on Management of Data*. New York: ACM, 2004: 623-634
- [45] Tao Yufei, Papadias D, Sun Jimeng. The TPR* -tree: An optimized spatio-temporal access method for predictive queries [C] // *Proc of the 29th Int Conf on Very Large Data Bases*. Trondheim, Norway: VLDB Endowment, 2003: 790-801
- [46] Zeitler E, Risch T. Massive scale-out of expensive continuous queries [C] // *Proc of the 37th Int Conf on Very Large Data Bases*. Trondheim, Norway: VLDB Endowment, 2011: 1181-1188
- [47] Yu Ziqiang, Liu Yang, Yu Xiaohui, et al. Scalable distributed processing of K nearest neighbor queries over moving objects [J]. *IEEE Trans on Knowledge and Data Engineering*, 2015, 27(5): 1383-1396
- [48] Cordeau J F. A branch-and-cut algorithm for the dial-a-ride problem [J]. *Operations Research*, 2006, 54(3): 573-586
- [49] Kalantari B, Hill A V, Arora S R. An algorithm for the traveling salesman problem with pickup and delivery customers [J]. *European Journal of Operational Research*, 1985, 22(3): 377-386
- [50] Herbawi W, Weber M. Ant colony vs genetic multiobjective route planning in dynamic multi-hop ridesharing [C] // *Proc of the 23rd IEEE Int Conf on Tools with Artificial Intelligence (ICTAI)*. Piscataway, NJ: IEEE, 2011: 282-288
- [51] Herbawi W, Weber M. The ridematching problem with time windows in dynamic ridesharing: A model and a genetic algorithm [C] // *Proc of IEEE World Congress on Computational Intelligence (WCCI 2012)*. Piscataway, NJ: IEEE, 2012: 1-8
- [52] Santos D O, Xavier E C. Dynamic taxi and ridesharing: A framework and heuristics for the optimization problem [C] // *Proc of the 23rd Int Joint Conf on Artificial Intelligence (IJCAI)*. San Francisco: Morgan Kaufmann, 2013: 2885-2891
- [53] Chaube V, Kavanaugh A L, Perez-Quinones M A. Leveraging social networks to embed trust in rideshare programs [C] // *Proc of the 43rd Hawaii Int Conf on System Sciences (HICSS)*. Piscataway, NJ: IEEE, 2010: 1-8
- [54] Cici B, Markopoulou A, Frías-Martínez E, et al. Quantifying the potential of ride-sharing using call description records [C] // *Proc of the 14th Workshop on Mobile Computing Systems and Applications*. New York: ACM, 2013: 17
- [55] Artz D, Gil Y. A survey of trust in computer science and the semantic Web [J]. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2007, 5(2): 58-71
- [56] Yu Han, Shen Zhiqi, Miao Chunyan, et al. A survey of trust and reputation management systems in wireless communications [J]. *Proceedings of the IEEE*, 2010, 98(10): 1755-1772
- [57] Witkowski J, Seuken S, Parkes D C. Incentive-compatible escrow mechanisms [C] // *Proc of the 25th AAAI Conf on Artificial Intelligence*. Menlo Park: AAAI, 2011: 751-757
- [58] Kamar E, Horvitz E. Collaboration and shared plans in the open world: Studies of ridesharing [C] // *Proc of the 21st Int Joint Conf on Artificial Intelligence (IJCAI)*. San Francisco: Morgan Kaufmann, 2009: 187-194
- [59] Makowski L, Ostroy J M. Vickrey-Clarke-Groves mechanisms and perfect competition [J]. *Journal of Economic Theory*, 1987, 42(2): 244-261
- [60] Cheng S F, Nguyen D T, Lau H C. Mechanisms for arranging ride sharing and fare splitting for last-mile travel demands [C] // *Proc of the 2014 Int Conf on Autonomous Agents and Multi-Agent Systems*. New York: ACM, 2014: 1505-1506
- [61] Zhao Dengji, Zhang Dongmo, Gerding E H, et al. Incentives in ridesharing with deficit control [C] // *Proc of the 2014 Int Conf on Autonomous Agents and Multi-Agent Systems*. New York: ACM, 2014: 1021-1028
- [62] Rayle L, Shaheen S, Chan N, et al. App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in San Francisco [R]. Berkeley: University of California Transportation Center, 2014: 1-19
- [63] Ozenc F K, Cranor L F, Morris J H. Adapt-a-ride: Understanding the dynamics of commuting preferences through an experience design framework [C] // *Proc of the 2011 Conf on Designing Pleasurable Products and Interfaces*. New York: ACM, 2011: 61
- [64] Mirisae S H, Brereton M, Roe P. Bridging the representation and interaction challenges of mobile context-aware computing: Designing agile ridesharing [C] // *Proc of the 23rd Australian Computer-Human Interaction Conf*. New York: ACM, 2011: 221-224
- [65] Radke K, Brereton M, Mirisae S, et al. Tensions in developing a secure collective information practice—the case of agile ridesharing [C] // *Proc of the 13th Int Conf on Human-Computer Interaction (INTERACT 2011)*. Berlin: Springer, 2011: 524-532
- [66] Fagnant D J. The future of fully automated vehicles: Opportunities for vehicle-and ride-sharing, with cost and emissions savings [D]. Austin: The University of Texas at Austin, 2014



Shen Bilong, born in 1984. PhD candidate. Student member of CCF. His main research interests include spatio-temporal data mining, urban computing, and machine learning.



Huang Yan, born in 1974. PhD supervisor. Her main research interests include spatio-temporal databases and mining, geo-stream data processing, smart transportation, and location-based social networks.



Zhao Ying, born in 1977. PhD. Member of CCF. Her main research interests include data mining, machine learning.



Zheng Weimin, born in 1946. PhD supervisor. Fellow member of CCF. His main research interests include grid and cluster computing, high performance storage system and big data analysis.

《计算机研究与发展》征订启事

《计算机研究与发展》(Journal of Computer Research and Development)是中国科学院计算技术研究所和中国计算机学会联合主办、科学出版社出版的学术性刊物,中国计算机学会会刊. 主要刊登计算机科学技术领域高水平的学术论文、最新科研成果和重大应用成果. 读者对象为从事计算机研究与开发的研究人员、工程技术人员、各大专院校计算机相关专业的师生以及高新企业研发人员等.

《计算机研究与发展》于1958年创刊,是我国第一个计算机刊物,现已成为我国计算机领域权威性的学术期刊之一. 并历次被评为我国计算机类核心期刊,多次被评为“中国百种杰出学术期刊”. 此外,还被《中国学术期刊文摘》、《中国科学引文索引》、“中国科学引文数据库”、“中国科技论文统计源数据库”、美国工程索引(EI)检索系统、日本《科学技术文献速报》、俄罗斯《文摘杂志》、英国《科学文摘》(SA)等国内外重要检索机构收录.

国内邮发代号:2-654;国外发行代号:M603

国内统一连续出版物号:CN11-1777/TP

国际标准连续出版物号:ISSN1000-1239

联系方式:

100190 北京中关村科学院南路6号《计算机研究与发展》编辑部

电话: +86(10)62620696(兼传真); +86(10)62600350

Email: crad@ict. ac. cn

<http://crad.ict. ac. cn>