

# 当出租车轨迹挖掘遇见智能交通

齐观德<sup>1</sup> 潘纲<sup>1</sup> 李石坚<sup>1</sup> 张大庆<sup>2</sup>

<sup>1</sup>浙江大学

<sup>2</sup>法国国立电信学院

关键词：智能交通 出租车轨迹

城市交通系统如同人体的血液循环系统，负责输送与配置城市的人力、物力资源，对城市的经济发展与市民生活有着重要影响。在中国城市化进程中，城市人口急剧增加，城市规模日益扩大，对交通资源数量和交通服务质量提出了极大挑战。传统的交通设施和交通服务已不能满足当前人们的交通需求，因而出现了交通堵塞、打车难等一系列问题。

信息技术的发展将会改善城市交通系统，使我们能够快速共享信息，合理调度交通资源，帮助我们建立智能交通系统，提升交通服务质量。泛在的传感设备采集城市交通数据，通过这些数据，我们可以了解到城市道路结构与道路通行状况、交通需求和供给的情况以及通行时间与路线导航信息。这些信息的集成将会推动城市交通系统的智能化改革<sup>[1]</sup>。

目前，许多城市的出租车上已经安装了GPS设备，产生着大量的轨迹数据。这些数据蕴含着城市交通系统的丰富信息，引

起了学术界的广泛关注，并催生了许多关于智能交通服务的研究<sup>[2]</sup>。研究表明，使用出租车轨迹数据挖掘可以帮助政府了解城市道路状况，交通资源与交通需求的分布，甚至道路与交通路线规划信息；可以帮助乘客了解出行的时间与费用以及交通服务质量；可以帮助司机推荐导航路线，改善出租车的运营策略。

## 轨迹数据的若干视角

简单地说，出租车轨迹数据是许多行采样记录，每一行包括当前时刻、车辆ID、经纬度、速度、载客状态(1/0表示有/无乘客)

和朝向信息(与正北方顺时针偏转的角度)。例如表1显示的是杭州市出租车轨迹数据的若干记录。由于采样设备的差异性，轨迹数据的采样频率不尽相同。以杭州市为例<sup>[3]</sup>，每辆出租车平均每分钟产生一条记录，每天可产生上千条记录，杭州现在有上万辆出租车，每天可产生千万条记录。

如何处理与分析这么多的数据记录？从数据处理的角度来看，可以把轨迹数据表示为GPS采样点的集合、二维地图上的曲线和关键点构成的字符串(如图1所示)。

**GPS采样点的集合** 在简单的情况下，考虑一个个离散

表1 出租车轨迹数据样例

时刻	车辆ID	经度	纬度	速度	朝向	载客状态
2009-04-01-00:00:04	1876	120.157295	30.241793	0	170	1
2009-04-01-00:00:04	14273	120.16118	30.272419	29.63	90	1
2009-04-01-00:00:04	2471	120.16782	30.284243	51.86	260	0
2009-04-01-00:00:04	14883	120.067444	30.090492	3.7	80	0
2009-04-01-00:00:07	18336	120.15485	30.290527	0.74	0	1
2009-04-01-00:00:07	10323	120.14411	30.327316	44.45	260	0
2009-04-01-00:00:14	16803	120.177246	30.292812	0	10	1
2009-04-01-00:00:21	16703	120.42536	30.232523	0	0	0

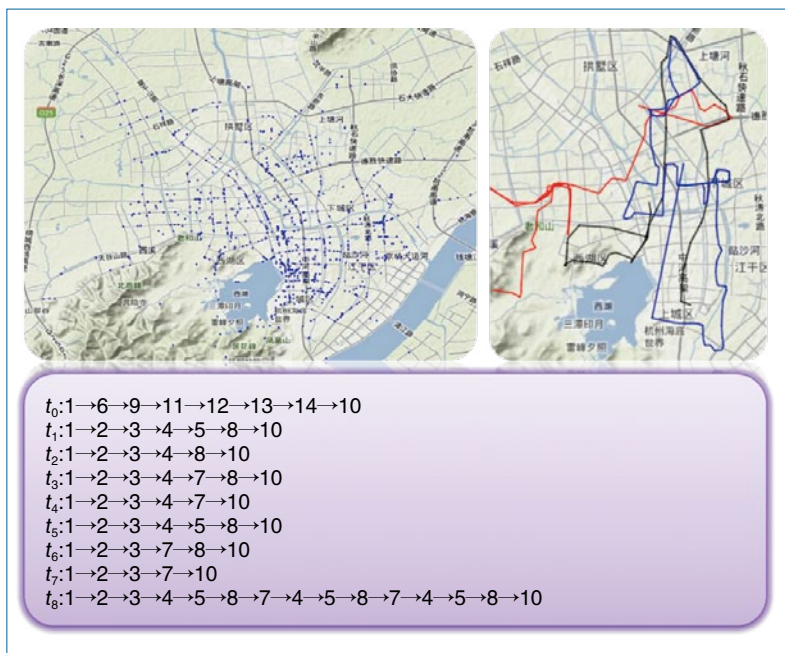


图1 轨迹的三种数据表现形式：左上为GPS点集合，右上为二维地图上的曲线，下方为关键位置组成的字符串

的GPS位置以及出租车在这些位置的行为。仅仅使用一些离散的GPS点，可以大致反映城市道路的轮廓<sup>[4]</sup>。而相应的速度信息，则反映了这些路段的交通情况<sup>[5]</sup>。此外，也要考虑一些特殊的GPS采样点，尤其是载客点/下客点<sup>[3]</sup>，即前一刻载客状态为1，当前载客状态为0的采样点。载客/下客地点反映了城市交通需求与人们的目的地。

**二维地图上的一些曲线**  
将每辆出租车的GPS采样按照时间先后顺序连接成一条曲线，就可以获得出租车的采样路径。采样路径和道路地图往往不重合，轨迹匹配算法<sup>[6]</sup>可以通过稀疏GPS采样推断出租车真正行驶的道路路径。生成的出租车行驶路径可用于轨迹聚类，识别出租车

的行为模式<sup>[7]</sup>，计算道路交通情况<sup>[5]</sup>，以及规划城市交通。

**关键地点构成的字符串**  
出租车的整条轨迹包含丰富的信息，但是在计算和表示上比较复杂

杂。根据二八定律，大部分路段只有很少的车辆经过，关注车辆频繁经过的少量路段或位置有助于简化计算。因此，轨迹可以表示成一些关键地点组成的字符串或者时间序列，使用字符串匹配可以比较不同轨迹<sup>[8]</sup>；使用概率图可以计算轨迹相关的概率，如搭载乘客的概率<sup>[9]</sup>。这些关键地点可以是车辆经常经过的路段<sup>[10]</sup>，也可以是一些热点区域<sup>[9]</sup>。当关键地点是热点区域时，还可以加入区域的语义信息，如区域社会功能。

从轨迹分析的角度，对于同一辆出租车的轨迹(如图2所示)，我们关注的内容可能也会不同。

**泛化的传感器** 忽略出租车本身的信息，而将出租车泛化为一个在道路上移动的传感器。此时，使用出租车轨迹数据，可以感知城市交通状况<sup>[5]</sup>，重建道路网结构<sup>[4]</sup>以及实时更新



图2 杭州市一辆出租车2009年4月1日的轨迹

路网信息。

**空载出租车轨迹** 只关注那些空载的出租车轨迹，观察出租车司机在空载时会采取哪些行为，以及这些行为是否能有效地帮助他们拉到乘客。通过研究可以为司机提供空载时寻找乘客的策略<sup>[11]</sup>，也可以为他们推荐最佳的驾驶路线<sup>[9]</sup>，包括时间最短、载客收益最大化、油耗最少的路线。

**载客出租车轨迹** 更多的研究是关注载客出租车的轨迹。出租车的载客路径反映了出租车司机的行为模式，其中一些行为可能是异常的（如绕路）<sup>[8]</sup>，一些可能有助于提高运营效率<sup>[12]</sup>，甚至可以用于道路导航<sup>[10]</sup>。出租车的载客路径还反映了人群（出租车乘客）可能的流动方向，可以用于分析道路交通流量或密度<sup>[13]</sup>，评估道路交通规划<sup>[14]</sup>。

**载客轨迹的端点** 载客轨迹的起始点与终止点对应乘客的出发地与目的地。通过挖掘出租车轨迹，可以计算并预测任意一对出发地/目的地之间需要花费的时间和费用，计算出发地/目的地对之间的乘客流量，并规划城市交通路线<sup>[15]</sup>。我们可以关注乘客的出发地，分析每个地方的乘客流量与空车流量，推荐合适的载客位置或者搭车位置<sup>[9]</sup>；也可以关注乘客的目的地，观察这些地方何时人流活动比较频繁，甚至识别这些地方的社会功能<sup>[3]</sup>。

## 政府：如何改善城市交通？

### 路网地图更新

**现状分析** 已经有多种使用出租车轨迹构建城市道路路网方面的研究，有的使用GPS位置<sup>[4]</sup>，有的使用GPS轨迹<sup>[16]</sup>。这些研究能识别与构建简单的道路形状、走向<sup>[4]</sup>，复杂的道路双向通行、多个车道<sup>[16]</sup>，甚至是三维的空间结构，如高架桥<sup>[17]</sup>。路网构建工作可以实时更新城市道路图，降低道路测绘工作的人力、物力成本。

**问题讨论** 目前路网构建算法的主要问题在于性能评估。由于缺乏实际的道路地图，一些研究将构建出的路网和OpenStreetMap<sup>1</sup>进行对比，但是OpenStreetMap是由共享的GPS数据构建的，不区分人行道与机动车道，而且也不包括一些数据中没有经过的道路。此外，研究一般通过道路的匹配程度作为性能的评价标准，而没有考虑道路拓扑结构是否得到保持。这样会出现一些严重问题，例如，如果仅仅丢失了道路的连接部分，匹配程度仍然会很高，但生成的地图却不能用于导航。

### 道路规划评估

**现状分析** 根据出租车轨迹与城市道路网，可以分析出租

车在道路上的移动情况，进而评估城市道路规划。微软亚洲研究院的郑宇等<sup>[14]</sup>提出了一种非常巧妙的方法，他们借助道路通行速度和道路弯曲程度来度量道路规划的好坏。道路通行速度越低，道路越绕，说明道路规划得越差。

**问题讨论** 使用出租车轨迹数据评估道路规划需要注意的一些问题是：(1) 出租车轨迹数据能否反映真实的交通情况？一些度量，如通行速度只依赖于道路状况，而不依赖于交通工具；而另一些依赖交通工具的度量，例如交通流量，则不能仅仅根据出租车轨迹数据推断。(2) 出租车轨迹是一种有偏轨迹，出租车行驶距离不一定是道路之间的最短距离。出租车司机对道路比较熟悉，会理性地选择合适的道路行驶，有时会根据路况选择绕路。因此，综合考虑交通速度和绕路是较为合理的评估标准。

### 交通供求分析

**现状分析** 如果不区分每次载客的人数，则可以简单地将载客轨迹的数量等价于乘客流量<sup>[18]</sup>。从而得出每个区域的乘客流量，即交通需求量，也可以将经过区域的空车数量作为交通供给量<sup>[19]</sup>。交通供求建模可以帮助政府了解城市交通需求与资源的匹配情况，重新调度并合理安排公共交通资源。

**问题讨论** 目前，这方面

<sup>1</sup> OpenStreetMap（简称OSM）是一个网上地图协作计划，目标是创造一个内容自由且能让所有人编辑的世界地图。

的工作仅限于空车与乘客流量的建模与预测，一般认为一定时间内到达的空车或乘客数量都是服从泊松分布的，并且这样的泊松分布会随着时间和地点的变化而变化。交通供求分析的主要问题在于：(1) 出租车的行驶路线不是固定的，交通供给量具有一定的随机性；(2) 乘客可以灵活自由地选择多种交通方式，很多乘客是潜在的出租车乘客，但是我们很难从出租车轨迹数据中发现这些潜在的乘车需求。

## 交通路线规划

**现状分析** 出租车乘客流很多时候只是城市公共交通人流的一个子集，不能直接作为城市人流的度量。在特殊情况下，例如凌晨，一些没有夜班公交的区域，出租车的乘客流量基本能反映当时、当地的人流。基于这样的想法，笔者提出了一个夜班公交线路算法，生成一条在尽可能短的时间内搭载尽可能多的乘客的公交线路<sup>[16]</sup>。实验结果表明，该算法生成的公交线路和白天的公交线路非常相似<sup>[15]</sup>（如图3所示）。

**问题讨论** 公共交通路线规划是一个复杂的问题，离实际应用还有距离。除了避免以偏概全，把出租车乘客流量作为公共交通客流量以外，我们还需要考虑：(1) 乘客流量在时空上会发生变化，一个合理的公交线路如何保证负载均衡，既不会在一些区域过载，也不会在一些时间段



图3 夜班公交线路规划

过载；(2) 提高载客效率，在提高载客量和缩短行车时间之间权衡；(3) 现实公交线路不止一条，采用多少条公交线路效率最高，如何交叉线路，如何分流客流量。

## 乘客：获得哪些智能服务？

### 交通时间分析

**现状分析** 狭义的交通时间是指人们乘坐交通工具从一地到另一地的通行时间，广义的交通时间还包括人们等候交通工具的到来所需要的时间（候车时间）<sup>[19]</sup>。通行时间的计算可以使用平均的两地通行时间<sup>[20]</sup>，也可以根据道路通行速度进行推断。候车时间的计算需要对车辆到达与人群等候队列进行建模，并利用排队论来求解。交通时间的计算有助于人们合理安排出行时间，规划行程。

**问题讨论** 使用出租车数

据计算交通时间具有一定的局限性。首先，出租车在两地的通行时间不等于其它交通工具的通行时间。公交车会在站点停靠，地铁显然更快，而私家车的通行时间也不同于出租车，因为出租车在很多时候会规避拥堵道路，选择与私家车不同的交通路线。其次，出租车的等候时间因为等候队列的灵活多变而难以计算，例如人们在等候出租车时并不会形成有序的等候队列，而且等候出租车的人群会因等候时间过长而离开。最后，要考虑天气、节假日等因素对道路交通的影响。

### 交通费用预测

**现状分析** 在日常生活中，如果我们频繁来往于两地之间，会很清楚两地之间的交通费用。如果能够获得出租车每条轨迹盈利的数据，那么预测交通费用也不是一件困难的事情<sup>[20]</sup>。然而，出租车轨迹数据通常没有盈利信息，我们需要通过分析出租车在

两地之间的频繁路线,综合道路长度、道路通行时间与红绿灯情况来预测交通费用。



图4 在安卓平台上开发的打的助手

**问题讨论** 出租车计价通常是与行驶距离、等待时间相关的一个二元函数,行驶距离在给定路线之后不会有太大变化,但等待时间会因为道路交通状况、红绿灯情况不同而差异较大。准确检测每个路口红绿灯是否存在,以及红绿灯的变化规律是非常困难的。一个粗略的方法是假设红绿灯在路网中等距分布,则推断等待红绿灯的时间与行驶距离成正比。

## 候车信息推荐

**现状分析** 我们经常会发现有些区域特别难打到出租车,而向周围行进一段距离就可以

了。通过对区域空载出租车数量和候车乘客数量建模,可预测乘客的等候时间(如图4)。乘客在



获悉等候时间后,可以自行选择是否等待。通过建模也可以为那些打车难的区域的乘客推荐更为合适的乘车地点<sup>[19]</sup>。

**问题讨论** 预测候车时间面临的困难有:(1)对于乘客等候队列的建模,出租车轨迹数据中并不直接包含乘客的到达时间,也不包含目前有多少名乘客在等候空载出租车的信息;(2)许多地方只有少量的乘客乘车,这些地方没有足够的支撑建模并预测候车时间。

候车地点的推荐则更为复杂,不仅要准确计算每个地点的等候时间以及乘客在地点间移动的时间,还要确定有多少

乘客会采用我们的推荐。推荐算法的一个悖论在于,如果很多乘客按照推荐前往容易打车的地点,则那些地点的乘客等候队列会增加,反而可能会变得不容易打车。

## 服务质量评估

**现状分析** 对于乘客而言,了解出租车司机的服务质量是非常重要的。一方面,出租车的服务质量包括乘车时间、费用、行车安全等诸多要素,这些要素的重要程度取决于乘客;另一方面,服务质量评估最终与出租车司机的驾驶行为相关。出租车的异常轨迹和其它大部分正常轨迹会有很大差别,通过比对可获取出租车司机的不恰当行为<sup>[8]</sup>。

**问题讨论** 服务质量评估的重要问题在于定义和识别出租车司机的语义行为。首先,轨迹数据的平均采样间隔一般在1分钟以上,以平均速度30km/h计算,出租车1分钟行驶距离可达500米,所以很难发现司机抢车道、强行超车这样细粒度的不安全行为,而只能定义一些全局的行为,如行车时间、费用、绕路等。其次,出租车的一些行为以某些标准衡量是好的,而以另一些标准衡量则可能是差的。例如,为了缩短行车时间,出租车司机可能会选择绕行,费用也有可能增加。最后,即使将出租车轨迹定义为正常或异常,也无法确定异常轨迹的详细语义,有经验或无经验的出租车司机表现得可能

和一般司机不一样。

## 司机：如何提高驾驶效率？

### 道路路况检测

**现状分析** 驾车人除了关心自己的车辆之外，还关心道路路况。道路通行、交通堵塞<sup>[21]</sup>与交通速度等情况都可以通过出租车数据挖掘发现<sup>[5]</sup>。

**问题讨论** 通过构建路网，根据路网的实时变化，检测道路通行状态，包括限行、单双向流通等。更简单的方法是在路网的基础上判断是否有车流经过。不管采用哪种方法，都要注意区分数据稀疏造成的路段信息缺失和真正的路段限行。

道路交通情况分析是交通领域的经典问题，包括道路通行速度与道路车流密度的计算。利用出租车轨迹数据来分析道路交通状况的问题有：(1) 出租车在道路上的密度以及它们对路线的选择不等于总的交通情况；(2) 出租车轨迹数据采样稀疏，每个路段上可能只有很少的数据来计算交通情况；(3) 如果根据出租车行驶路程与时间的比值来计算交通速度，则需要排除等待红绿灯所耗费的时间。

### 行车路线推荐

**现状分析** 城市交通路网交错，任意两地之间存在多种可行路线。交通情况多变，道路拥

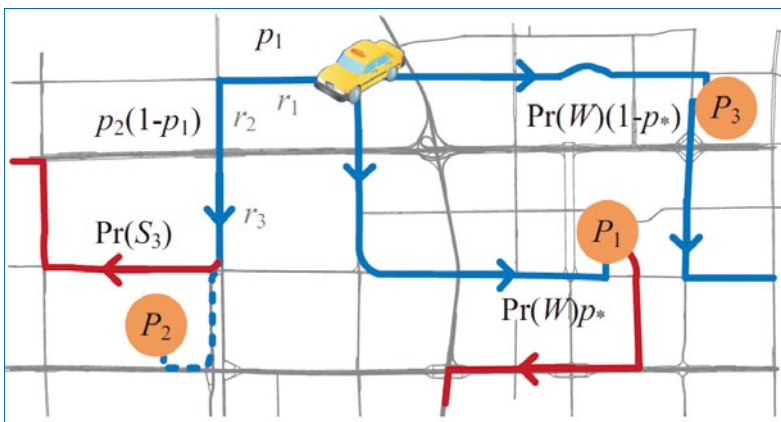


图5 计算沿途搭载乘客的概率推荐载客路线<sup>[9]</sup>

塞程度因时因地而异。城市交通导航最重要的不是选择最短路径，而是避开拥塞路段，选择最快的路径。出租车司机具有丰富的驾驶经验，熟悉城市道路交通情况，因而往往能够选择最有效率的路线。通过挖掘出租车轨迹数据，可以学习出租车司机的智慧，获得更好的行车路线<sup>[10]</sup>。

**问题讨论** 行车路线推荐面临的主要问题在于它不可避免地遇到推荐的悖论，即如果多数人都按照推荐的路线行车，则推荐的路线反而变得拥堵。真正理想的城市交通应该是负载均衡的，推荐行车路线的目的是最终使每条路线上的交通时间相近。此问题从计算的角度必然是非确定性多项式难题(NP-hard)。可行的方法是利用群体智慧，即不仅考虑当前的路况，还考虑人们对路线的选择，让每个驾驶者都知道每条路线选择人数的多少，让驾驶者自行博弈、选择路线。

### 载客策略推荐

**现状分析** 载客策略推荐是帮助出租车司机更快、更好地找到乘客的一种方法。载客策略包括载客地点的选择<sup>[18]</sup>、载客路线的推荐<sup>[9]</sup>（如图5所示）以及寻找乘客的方式<sup>[11]</sup>。载客地点选择的目的在于告诉出租车司机当前去什么位置更容易载客。载客路线在载客地点的基础上考虑沿途搭载乘客的概率，推荐最优的载客路线，使得载到乘客的概率最大化。除了按照一定的路线主动寻找乘客外，还有一种载客策略是在合适的地方等待。

**问题讨论** 载客策略推荐并不是一个简单的问题，除了要考虑推荐会引起出租车司机之间的博弈外，还要考虑：(1) 沿着最佳推荐路线的油耗与原地等待耗费的时间；(2) 有的区域可能更容易载到长途客人，从而给司机带来更大收益，但是这种推荐是否对短途乘客不公平？

### 收益评估

**现状分析** 收益评估是指

通过建立出租车驾驶行为模式与收益的关系,预测驾驶行为可能带来的收益,并推荐更好的驾驶行为。美国麻省理工学院 Senseable City 实验室的刘亮(Liu Liang,音译)等<sup>[12]</sup>发现出租车司机的收入和他们的驾驶策略有很大关系,收入高的司机具有相似的活动模式,他们在交通高峰期避开闹市拥堵地段,在交通低谷时进入闹市。

**问题讨论** 影响出租车司机收益的因素有很多,主要包括:(1)载客距离:根据出租车的计价方式,当载客距离超过一定范围时,出租车司机将获得更高的收益/距离比;(2)载客时间:除了偏好远距离的乘客,出租车司机减少空载时找乘客的时间有助于提高收入;(3)单位时间有效收益:城市交通情况越好,出租车单位时间行驶路程越多,收益效率越高。一般认为,城市白天的人群流动大,交通需求多,出租车司机基本处于载客运营状态,道路交通状况是影响其收入的主要因素;到了晚上,尤其凌晨,道路交通状况非常好,交通需求集中在少数夜市区域,出租车司机会耗费更多时间等待或寻找乘客。

## 结语

出租车轨迹数据是一种广泛存在、容易获取的公共交通数据,它包含了城市人群移动与城市交通的宏观信息。通过挖掘出租车

数据,研究者提出了许多改善现有交通系统的算法与相应的服务。这些理论与应用已经广泛应用于城市交通系统中,并将显著改善城市交通服务质量。但是,交通系统是一个复杂系统,出租车数据只能反映交通系统的一个侧面,在很多时候具有一定的局限性。我们希望随着传感技术的发展与算法理论的进步,可以获得全面的交通数据,完善智能交通系统,改善城市生活。■



**齐观德**

CCF会员。浙江大学博士生。主要研究方向为普适计算、移动数据挖掘等。  
qiguande@zju.edu.cn



**潘纲**

CCF高级会员、本刊编委。浙江大学教授。主要研究方向为普适计算、计算机视觉等。  
gpan@zju.edu.cn



**李石坚**

CCF会员。浙江大学副教授。主要研究方向为普适计算、移动互联网、社会计算等。  
shijianli@zju.edu.cn



**张大庆**

CCF会员。法国国立电信学院教授。主要研究方向为普适计算等。  
Daqing.Zhang@it-sudparis.eu

## 参考文献

- [1] Pan, G. and Qi, G. et al., Trace analysis and mining for smart cities: issues, methods, and applications, IEEE Communications Magazine, 51(6):2013, 120~126
- [2] P.S. Castro, D. Zhang, C. Chen, S. Li and G. Pan, From taxi GPS traces to social and community dynamics: a survey, ACM Computing Surveys, 2013, accepted
- [3] Pan, G. and Qi, G. et al., Land-use classification using taxi GPS traces, IEEE Transactions on Intelligent Transportation Systems, 14(1):2013, 113~123
- [4] Davics, J. J. and Beresford, A. R. and Hopper, A., Scalable, distributed, real-time map generation, IEEE Pervasive Computing, 5(4):2006, 47~54
- [5] Liu, S. et al., Towards mobility-based clustering, Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010, 919~928
- [6] Lou, Y. et al., Map-matching for low-sampling-rate GPS trajectories, Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2009, 352~361
- [7] Zhang, D. et. al., iBAT: detecting anomalous taxi trajectories from GPS traces, Proceedings of the 13th international conference on Ubiquitous computing, 2011, 99~108
- [8] Zheng, K. and Zheng, Y. et al., On discovery of gathering



## 2013年度CCF会士提名工作启动

根据《中国计算机学会会士条例》，2013年度CCF会士提名工作现已启动。CCF会士和CCF理事具有提名权，提名者须在2013年9月30日前将“CCF会士提名表”填妥后发送至 awards@ccf.org.cn。学会已组成以陈左宁副理事长为主席的工作委员会专事此项工作。

### 被提名 CCF 会士候选人的资格

- ▶ 在提名截止时必须具有 CCF 高级会员及以上会员资格
- ▶ 在提名截止时至少有 5 年 CCF 连续会龄（不包括学生会会员时期的会龄）

### 提名人的资格

- ▶ CCF 会士或 CCF 理事
- ▶ 提名人在提名截止时必须具有 CCF 高级会员及以上会员资格

### 提名要求

- ▶ 每位提名人本年度提名的 CCF 会士候选人数不得超过 2 人
- ▶ 提名人须保证所提名的会士候选人获得另外 3~5 名 CCF 会士的推荐

详见 [www.ccf.org.cn](http://www.ccf.org.cn)

patterns from trajectories, IEEE International Conference on Data Engineering, 2013

[9] Yuan, J. and Zheng, Y. et al., Where to find my next passenger, in Proceedings of the 13th international conference on Ubiquitous computing, Beijing, China, September 2011, 109~118

[10] Yuan, J. and Zheng, Y. et al., T-drive: driving directions based on taxi trajectories, Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2010, 99~108

[11] Li, B. and Zhang, D. et al., Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset, IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2011, 63~68

[12] Liu, L. and Andris, C. and Ratti, C., Uncovering cabdrivers' behavior patterns from their digital traces, Computers, Environment and Urban Systems, 34(6), 2010, 541~548

[13] Castro, P. and Zhang, D. and Li, S., Urban traffic modelling and prediction using large scale taxi GPS traces, International Conference on Pervasive Computing, 2012, 57~72

[14] Zheng, Y. et al., Urban computing with taxicabs, Proceedings of the 13th International Conference on Ubiquitous Computing, 2011, 89~98

[15] Chen, C. and Zhang, D. et al., B-Planner night bus route planning using large-scale taxi GPS traces, 11th IEEE Int. Conf. on Pervasive Computing and Communications, 2013

[16] Cao, L. and Krumm, J., From GPS traces to a routable road map, Proceedings of the International Conference on Advances in Geographic Information Systems, 2009, 3~12

[17] Fathi, A. and Krumm, J., Detecting road intersections from GPS traces, Geographic Information Science, 2010, 56~69

[18] Li, X. and Pan, G., Prediction of urban human mobility using large-scale taxi traces and its applications, Frontiers of Computer Science, 6(1):2012, 111~121

[19] Zheng, X. and Liang, X. and Xu, K., Where to wait for a taxi? Proceedings of the ACM SIGKDD International Workshop on Urban Computing, 2012, 149~156

[20] Balan, R. K. et al., Real-time trip information service for a large taxi fleet, Proceedings of the 9th international conference on Mobile systems, applications, and services, 2011, 99~112

[21] Peng, C. et al., Collective human mobility pattern from taxi trips in urban area, PloS one, 7(4):2012, e34487