

城市交通的可视分析研究

王祖超 郭翰琦 袁晓如
北京大学

关键词：交通轨迹可视分析 微博可视分析

城市化促进了经济的发展，改善了人们的生活水平，但同时也带来了严重的交通问题。为了缓解交通压力，人们开始利用传感器技术来解决问题，由此产生了各种各样的交通数据，包括车辆的速度、密度，道路占有率，交通视频，车辆和行人轨迹，以及有关交通的新闻和微博。在大数据时代，获取和存储这些数据变得越来越容易。然而，如何分析和利用这些海量又复杂的数据成为一项挑战。

可视分析技术为我们提供了一种直观有效的方法。它将复杂的交通数据及其分析结果通过可视化方式直观地展现出来，并支持对结果的交互式筛选和浏览。此外，

当人们从可视化产生的图像中发现一些意料中或者意料外的特征时，也可以启动分析算法来有针对性地自动深入挖掘交通信息。

本文主要讨论交通轨迹数据和微博数据的可视分析。根据我们最新的一些研究成果，文章从一个路口的交通分析出发，逐渐扩展到整个城市，包括城市交通密度的绘制，交通轨迹的交互筛选，以及交通拥堵的可视分析，最后介绍微博上的交通信息挖掘。

路口交通的可视分析

路口是城市交通的基本组成单元。由于信号灯的调节，车流

与车流、车流与人流的交汇，使得路口的交通状况十分复杂。为了研究路口交通，我们使用6台激光扫描仪，在一个路口采集了两天的交通轨迹数据，包括来往的机动车、自行车和行人，数据大小为4.7GB，包括209426条轨迹、33362651个采样点，每26毫秒采样一次。我们希望能够从宏观上展示路口交通的基本时空特性，从微观上寻找一些异常事件，例如交通事故和违规。为此我们开发了TripVista系统^[1]。该系统从空间、时间和属性三个角度对交通轨迹数据进行可视分析。

在空间上，我们将不同类型的移动物体的轨迹用不同的颜色

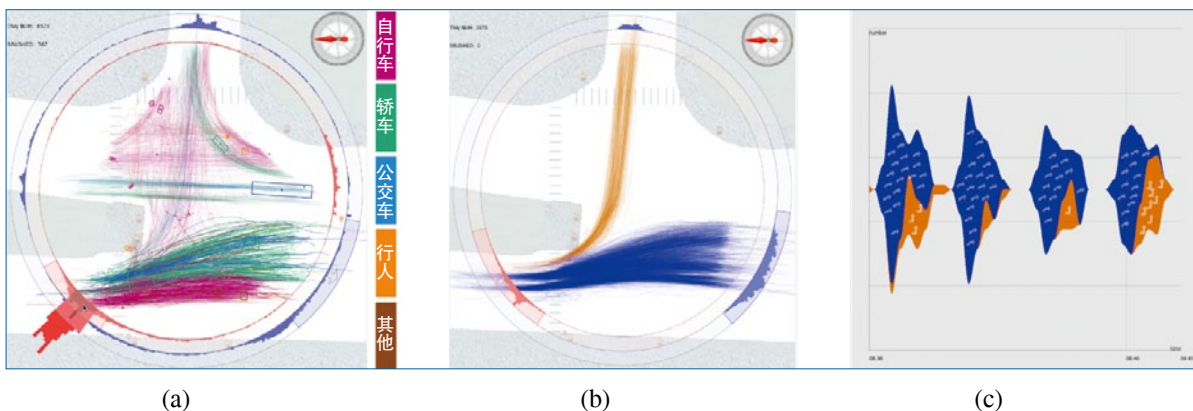


图1 路口交通总体规律的分析：(a)各种类型的移动物体在路口的空间分布；(b)根据物体进入和离开路口的方向选择两束轨迹；(c)使用主题河技术观察(b)中两束轨迹交通流量随时间的变化

绘制出来。如图 1(a)，它们的空间分布各不相同。图中两个大圆环上显示的是轨迹进入（红色）和离开（蓝色）路口的角度分布直方图，右上角的指南针指向北方。灵活的交互可以帮助用户选择从特定方向进入和离开的轨迹。如图 1(b)，从左下方路口进入的交通流被选中，并按离开路口的方向分成了两部分，分别用蓝色和棕色表示。

在时间上，我们通过一种叫做主题河的方式来展现交通流量随时间的变化。由此可以观察到早晚高峰，还可以看到路口信号灯调节对于交通流量的影响，如图 1(c)，其中横轴表示时间，两种颜色分别对应图 1(b) 中的两束交通流，每种颜色的纵向宽度表示相应交通流在该时刻的流量，内部的白色小箭头提示交通流的方向。我们可以看到流量随交通信号灯调节的周期性变化，周期约为 3 分钟。这种分析能对红绿灯参数的设置提供一定的参考。

在属性上，我们通过平行坐标的方式展现轨迹不同属性间的相关性。这些属性包括数据本身记录的起始时间、移动物体类型，也包括计算得到的物体通过路口的时间、最小速度、拐弯角度变化量总和等等。

对交通轨迹数据从时间、空间、属性三个方面的分析可以帮助我们大量的轨迹中筛选出一些异常行为，如车辆违章和交通险情，如图 2。图 2(a) 中一辆小轿车既不愿意等红灯，又不想闯

红灯，于是绕行了上方的单行道；图 2(b) 中一辆从右下角出发斜穿马路的自行车，几乎被两辆小轿车撞上。发现和分析这些行为能让道路交通更安全。

城市交通的可视分析

对于整个城市的交通，人们关心的问题包括城市的交通密度、出发地/目的地以及一些特殊的交通事件（例如堵车）。为此我们开展了大量的研究工

作。我们使用北京市 24 天的出租车 GPS 轨迹数据，该数据大小为 34.5GB，包括 28519 辆车、379107927 个采样点，每 30 秒采样一次（但有大量采样点缺失）。该数据包含了北京市 43% 的出租车，大约对应北京市四环以内 7% 的交通流量，因此该数据可以大致反映北京市的整体交通状况。

对于城市的交通密度，我们使用密度图来显示，它反映了各区域交通的繁忙程度。图 3 为北

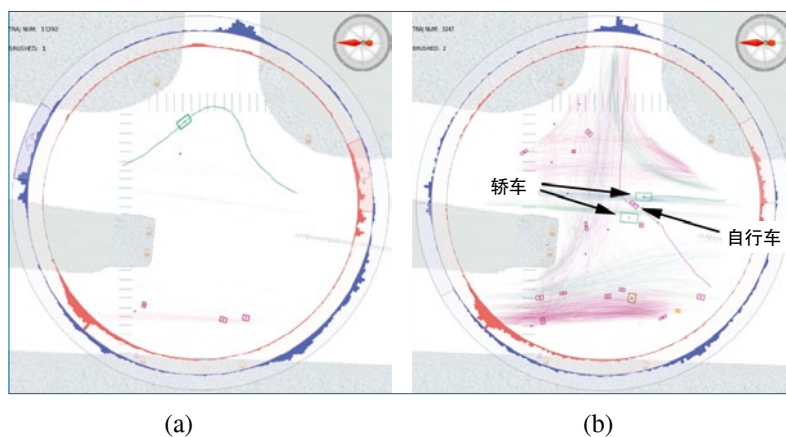


图2 路口交通异常事件的挖掘：(a)从右往左交通流中的异常轨迹；(b)交通险情

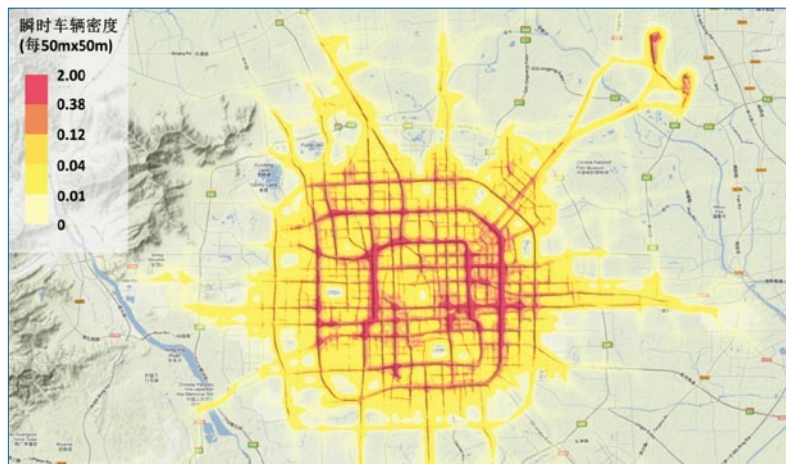


图3 北京市一天出租车轨迹密度图（颜色表示瞬时车辆密度，即任意一个瞬间在单位区域内观察到车辆数目的平均值）



图4 基于环形选择器的轨迹筛选

北京市 2009 年 3 月 8 日的出租车轨迹密度图。

对于出发地 / 目的地分析, 我们希望能够从大量轨迹中方便地筛选出具有特定出发地 / 目的地的轨迹, 以便比较它们的时间分布、属性区别和路径选择。为此, 我们设计了一些环形的筛选器, 用于指定轨迹的起点、终点、经过点、时间、属性等条件, 从而提取出符合条件的轨迹。用户只须进行简单的鼠标操作, 而无须编写复杂的查询语句。图 4 展示的是筛选结果, 以左下和右上两个环形区域为起止点, 并且途径中间环形区域的载客出租车轨迹被筛选出来。环外的黄圈上显示了轨迹经过环形区域的时间统计信息。

在交通事件方面, 交通拥堵是一个复杂问题。拥堵本身会随着路段和时间的不同而变化, 此外它还会沿着路网传播。我们工作的重点在于帮助人们理解和探索这些复杂因素。为此我们建立了一个城市交通拥堵的可视分析系统^[2]。我们基于北京市 24 天的出租车轨迹数据, 提取出了一系列道路交通拥堵事件, 并分析了这些拥堵在路网上的传播方式。

在单条道路层面, 采用图 5 的形式来表示每条道路 24 天的交通状况, 其呈现方式如同一个表格。其中 24 行表示 24 天, 144 列表示一天的 144 个 10 分钟。每个单元格表示某一天的某个 10 分钟时间段, 其颜色表示道路上

交通流的平均速度。大的单元格表示此时道路拥堵。

图 6 选择的是北京市 7 条有代表性的道路。大部分道路都会在工作日的早晚高峰和下午拥堵, 如 (a) 北三环中路主路马甸桥段。有些道路会受到周围单位的影响, 如 (b) 北四环西路辅路中关村段, 它临近中关村第一小学和中关村第三小学, 因为家长接送孩子上学放学, 拥堵时间为工作日的早上 7 点半之前和下午 4 点半之后的一段时间。同一条道路的相反方向拥堵时间可能差别很大, 例如莲花池东路主路北京西站段, 入城方向 (c) 早高峰拥堵, 出城方向 (d) 晚高峰拥堵。中国国际展览中心顺义新馆南侧的天北路 (e), 其拥堵日期与展览日期吻合。机场高速去往市区方向 (f) 的大部分时间通畅, 但偶尔会有拥堵, 无明显的规律。工人体育馆西路 (g) 附近有大量酒吧, 晚上比白天拥堵, 尤其是周五周六的晚上。图中地图上道路的颜色表示每天的平均堵车时间, 由黄到红逐渐变长。

我们可以使用轨迹动画来验证自动检测得到的拥堵事件。图 7 显示了一次拥堵事件 (对应图 6(f) 中的红色单元格) 动画的关

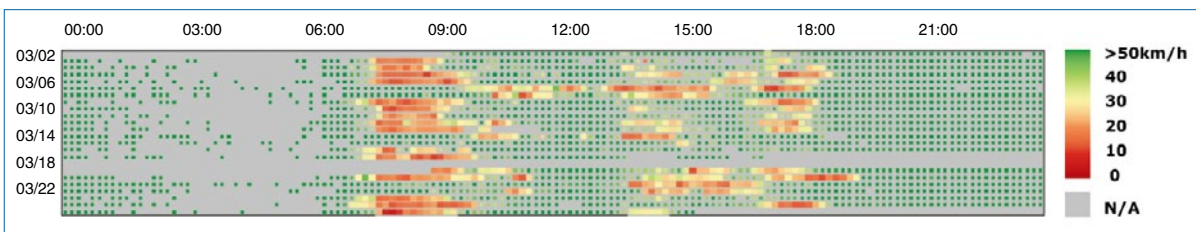


图5 单条道路交通状况的可视化

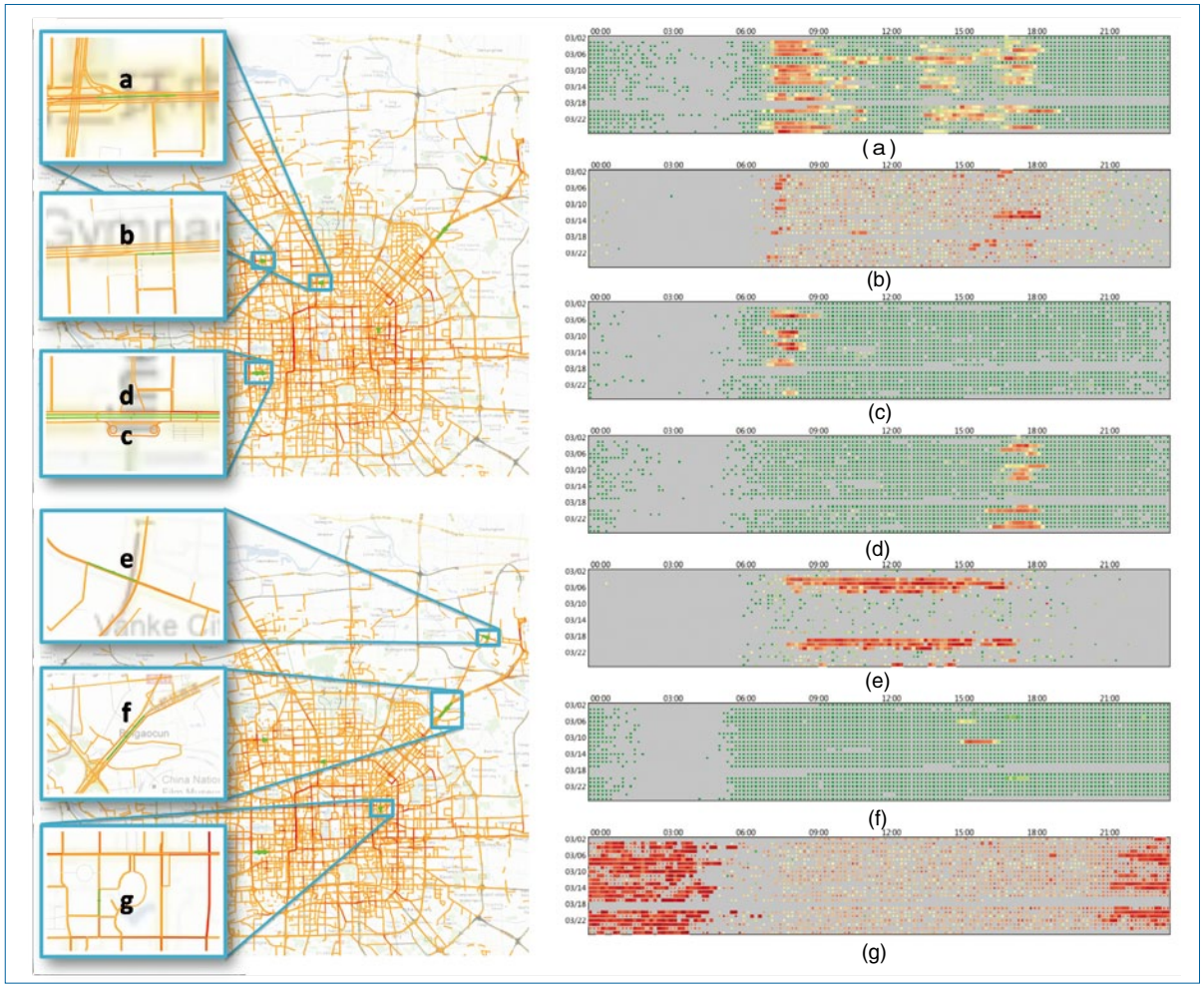


图6 7条代表性道路的交通拥堵情况

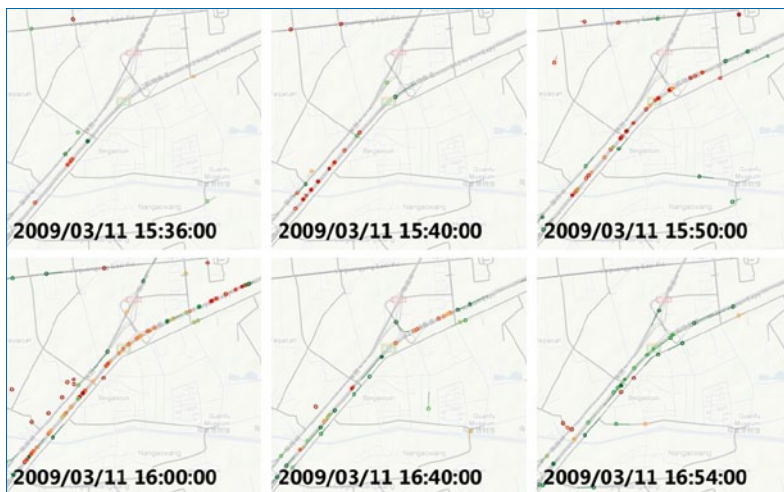


图7 使用轨迹动画观察自动检测得到的拥堵事件

键帧，其中每一个小圆点表示一辆出租车的位置。颜色表示速度（定义同图5）。此次拥堵发生在机场高速从机场到市区方向的路段，动画显示拥堵从15:36开始，到16:54结束，与自动检测结果（15:30~17:00）相符。

在单独分析每条道路之后，将相邻道路的拥堵联系起来，计算拥堵在道路上的传播过程。图8(a)显示的是2009年3月21日在北京市健翔桥发生的一次拥堵的传播路径。它有两个起点D和H，三个终点F、G和L。拥堵的

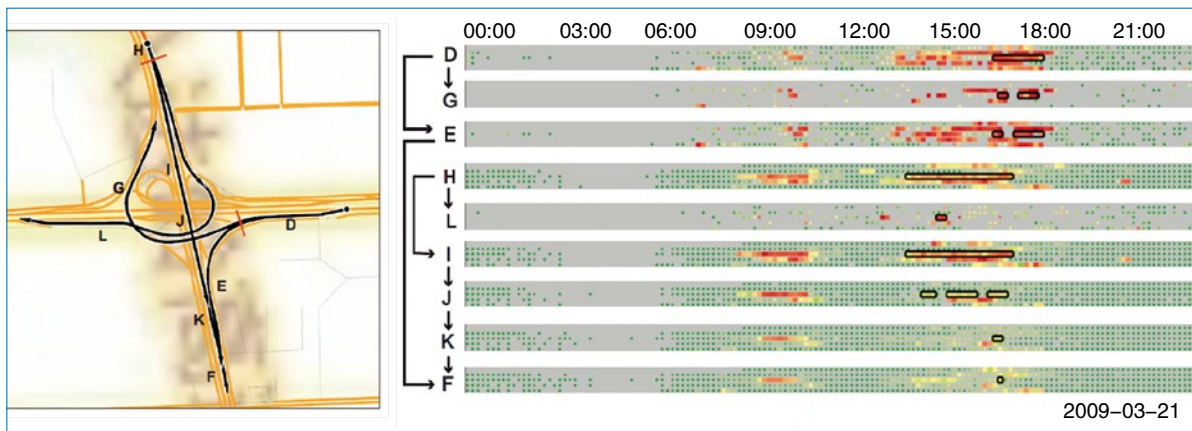


图8 健翔桥附近的一次拥堵传播: (a)传播路径; (b)拥堵道路的速度随时间的变化

传播时间从 14:10 开始, 到 18:40 结束, 共持续 4 小时 30 分钟。图 8(b) 显示了此次拥堵中每条道路上 (D~K) 交通流的速度随时间的变化, 其中每一个小图表示一条道路的速度变化, 黑框表示该条道路在此次拥堵传播过程中拥堵的时间。

午拥堵传播的变化情况。如果某天上午有多次独立的拥堵时间, 则选择持续时间最长的一次拥堵; 如果某天上午没有拥堵, 则不显示。图中箭头表示传播方向, 道路的颜色表示该道路的拥堵时间, 从黄色 (10 分钟) 到暗红色 (大于 4 小时) 拥堵时间越来越长。

有细微的不同。

微博上的交通信息挖掘

近年来, 微博成为人们获取信息的重要渠道。相比门户网站, 微博具有时效性高、公众参与性强、覆盖范围广等特点。社会生活各个角落的故事都会被写成微博发布出来, 这使得微博成为一种社会传感器。基于此, 我们开发了 ThemeMap 应用^[3]。它能够围绕指定的话题自动从新浪微博收集相关信息, 并通过用户标注的方法, 更正其中的错误, 补充不完整的信息, 最终构建出关于此话题的主题地图。

我们将 ThemeMap 用于交通信息的分析中, 构建了全国高速公路堵车、国庆北京堵车、厦门堵车、北京暴雨、北京打车难、北京堵车等主题地图。图 10 展示了“厦门堵车”话题的主题地图, 其中每一个小圆圈表示一条微博, 背景从蓝到红的颜色渐变表示微博的密度 (红色区域密度

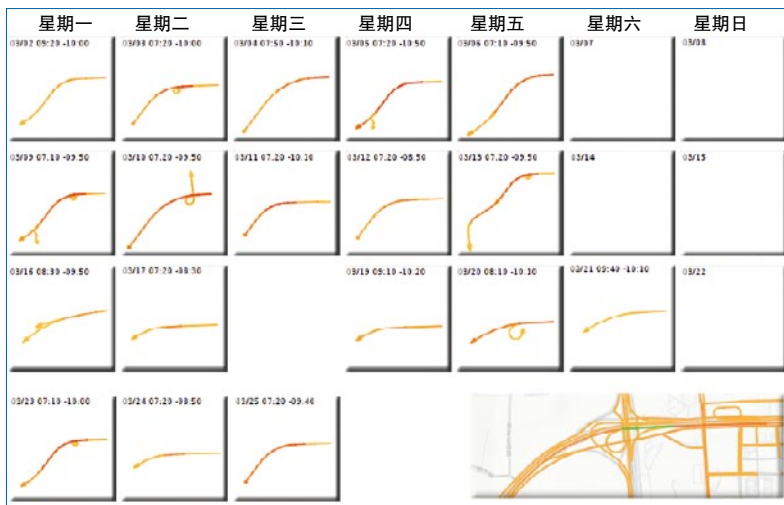


图9 万泉河桥24天内每天上午拥堵传播的路径比较

同一块区域, 其拥堵的传播方式会随时间变化。图 9 展示了北京市万泉河桥 24 天内每天上

由图中可以看出, 万泉河桥主要在工作日拥堵, 而在周末则很少拥堵。另外, 每天的拥堵模式都

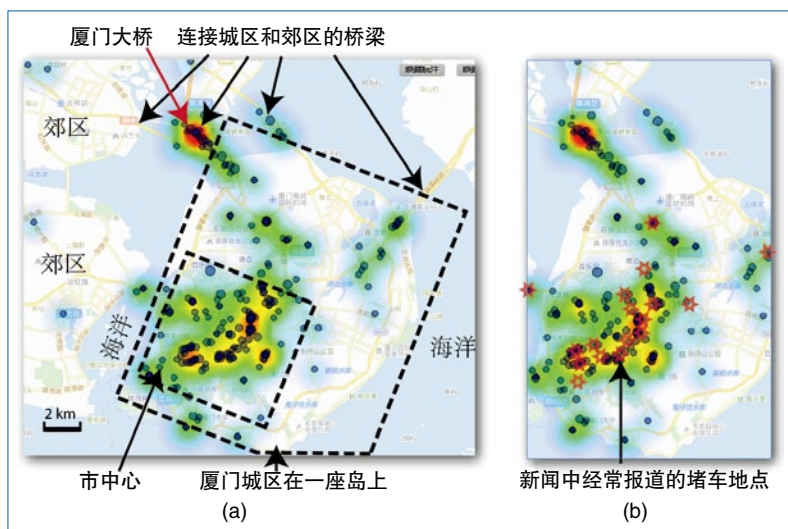


图10 (a)ThemeMap得到的“厦门堵车”主题地图；(b)结果与新闻中经常报道的堵车地点（五角星）相符

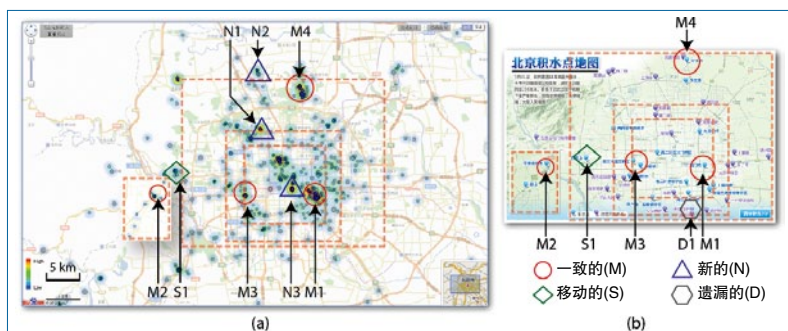


图11 (a)ThemeMap得到的“北京暴雨”主题地图；(b)结果与腾讯网“北京积水点地图”基本吻合

高)。它是在厦门一位交警的提议下创建，并由网友标注得到的。2012年7月21日，北京遭遇暴雨袭击，造成重大伤亡和经济损失。图11展示了“北京暴雨”话题的主题地图。该地图显示了受灾地区的分布和各地区遭受暴雨灾害的严重程度，为人们今后的出行安全提供参考。

交通数据的可视分析是一个新兴领域。我们相信针对交通数据的可视分析有着广阔的应用前景，能够帮助城市交通更加安全

顺畅地运行。■



王祖超

CCF学生会会员。北京大学博士生。主要研究方向为时空轨迹数据的可视分析等。zuchao.wang@gmail.com



郭翰琦

CCF学生会会员。北京大学博士生。主要研究方向为大规模数据可视化等。guohanqi@gmail.com



袁晓如

CCF高级会员、杰出演讲者、本刊编委。北京大学研究员。主要研究方向大规模数据高性能绘制和可视化等。xiaoruyuan@pku.edu.cn

参考文献

- [1] Hanqi Guo, Zuchao Wang, Bowen Yu Huijing Zhao and Xiaoru Yuan. TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In Proc. IEEE PacificVis, 163~170, 2011
- [2] Zuchao Wang, Min Lu, Xiaoru Yuan, Junping Zhang and Huub van de Wetering, Visual traffic jam analysis based on trajectory data. IEEE Trans. Vis. Comput. Graph. (VAST'13), 19(12), 2013. (ACCEPTED)
- [3] ThemeMap, <http://vis.pku.edu.cn/weibova/thememap>