

OD 数据可视分析综述

涂乐¹⁾, 陈彬捷¹⁾, 周志光^{1,2)*}

¹⁾ (浙江财经大学信息管理与人工智能学院 杭州 310018)

²⁾ (浙江大学 CAD&CG 国家重点实验室 杭州 310058)

(zhgzhou1983@163.com)

摘要: OD 数据是一类提供起点和终点位置、稀疏地描述对象移动轨迹的数据. OD 数据可视分析能够发现群体移动模式, 挖掘空间异常和隐藏关系, 深入探索和分析多种统计属性, 是数据可视化、地理空间分析等领域的研究热点. 从 OD 数据的地理空间点对结构出发, 阐述经典的 OD 数据可视化方法和原理, 如 OD 矩阵、OD 嵌套地图、OD 流图等; 然后从空间聚类维度、空间语义维度和时空联合维度对 OD 数据进行可视分析, 探寻其在特征的表达、感知、提取、呈现等方面的应用场景和增强意义; 再针对大规模 OD 数据可视化的视觉混淆问题, 梳理 OD 数据可视分析研究方法中过滤、采样及聚合等视觉简化方法; 最后通过 OD 数据可视分析技术在不同行业领域的应用及具体案例, 总结和深入讨论 OD 数据可视分析中存在的挑战性问题, 并展望该领域的发展趋势.

关键词: OD 数据; 可视分析; 视觉简化; 可视化应用

中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18679

A Survey on OD Data Visualization

Tu Le¹⁾, Chen Binjie¹⁾, and Zhou Zhiguang^{1,2)*}

¹⁾ (School of Information Management and Artificial Intelligence, Zhejiang University of Finance and Economics, Hangzhou 310018)

²⁾ (State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310058)

Abstract: OD data is a type of data that provides origin and destination positions, and sparsely describes the movement trajectory of objects. Being the research hotspots in data visualization and geospatial analysis, visual analysis of OD data contributes to discovering group movement patterns, exploring spatial anomalies and hidden relationships, and in-depth analysis of various statistics attributes. It is started from the pair-point structure of OD data in geospatial space, and classical OD data visualization methods and principles are elaborated, such as OD matrix, OD map, OD flow, etc. Then from the perspective of spatial-temporal feature exploration, spatial and spatial-temporal joint dimensions, the visual analysis of OD data is performed, while its application scenarios and enhanced significance of feature expression, perception, extraction and presentation are researched. To solve the problem of visual clutter in large-scale OD data visualization, the visual simplification methods such as filtering, aggregation, and sampling are sorted out in the research methods of visual analysis of OD data. Finally, the applications and specific cases of OD data visual analysis technology are investigated in different domains, the challenging problems in the research of OD data visual analysis are summarized, and the trend of its future development is prospected.

收稿日期: 2020-09-25; 修回日期: 2021-03-23. 基金项目: 国家自然科学基金(61802339, 61872314); 浙江省统计重点研究项目(20TJZZ10); 浙江省高校重大人文社科攻关计划(2018QN021); 浙江省科技厅公益项目(LGF20F020010); 浙江大学 CAD&CG 国家重点实验室开放课题(A2001). 涂乐(1989—), 男, 博士, 讲师, 主要研究方向为数据可视化与可视分析; 陈彬捷(2001—), 女, 在校学生, 主要研究方向为数据可视化与可视分析; 周志光(1983—), 男, 博士, 教授, CCF 会员, 论文通讯作者, 主要研究方向为数据可视化与可视分析.

Key words: origin destination data; visual analysis; visual abstraction; visualization application

随着位置感知技术的不断发展和大量应用,越来越多的轨迹数据被实时记录和获取,通常包含移动对象(如人类、车辆、动物等)、移动起点和终点、移动过程轨迹以及各个空间位置对应的时间信息.与轨迹数据相关的研究在应急疏散、交通运输、城市规划、传染病防控等领域具有重要意义^[1].为直观地呈现轨迹数据以便理解,国内外学者在轨迹数据可视分析方面做了大量的研究工作.Andrienko 等^[2]将轨迹数据可视化方法划分为直接可视化、聚集可视化和特征可视化.蒲剑苏等^[3]在轨迹数据的重建校准、空间和时间属性的可视化、用户交互系统等方面进行综述介绍.王祖超等^[4]更加详尽地介绍各类轨迹数据可视分析方法,列举应用案例并阐明相关研究面临的主要问题与挑战.在前人轨迹数据可视分析综述的基础上,本文对当前OD (origin destination)数据可视分析进行全面的调研与探讨.通过在 Web of Science、中国知网和万方等中英文数据库,以“Origin Destination Data Visualization”“OD 数据可视化”为主题进行检索,将会议论文、书目章节及不相关论文等剔除后,共获得英文文献 52 篇,中文文献 48 篇.本文将对这些中英文文献进行分析、总结和归纳.

OD 数据是一类稀疏的轨迹描述数据,主要包含空间迁移的起点和终点,仅能部分表达对象的移动轨迹,如出租车的上下客地理位置、手机基站记录用户的移动位置变化、用户在不同地点发布的社交媒体数据等.OD 数据特点如下.

(1) 以 OD 点对(起点、终点以及对应时间等属性)的形式存储,与空间地图结合紧密^[5].OD 数据由系列 OD 点对构成,通常包含空间位置及其流向关系,不同的空间位置和时间记录可以合成地理网络关系数据.因此,如何有效地呈现 OD 数据及其在地图上的分布关系,是 OD 数据可视化研究的基础.

(2) OD 数据中具有时空多维属性,节点、边的关系复杂,简单的 OD 可视化方法难以发现复杂特征^[6].为方便用户理解 OD 数据特征并解决相关领域问题,大量学者引入空间分析和时空联合分析模型和方法,以有效地增强感兴趣特征的表达、感知、提取和呈现.

(3) OD 数据量大,容易引起视觉遮挡和混淆^[7].由于传感器技术的快速发展,OD 数据采集多样化、

高频化和细致化,使 OD 数据的规模快速累积,进而大规模 OD 数据可视化将出现节点、连线大量重叠、交叉和覆盖等问题.因此,大量学者通过研究简化表达的方法来降低视觉遮挡,增强视觉感知.

(4) OD 数据类型普遍存在于各个应用领域,面向不同领域的 OD 数据可视分析和挖掘存在不同的交互设计规则和个性化的任务需求^[8].因此,大量研究工作从数据和应用的角度出发,针对 OD 数据进行定制的可视化分析探索和研究.

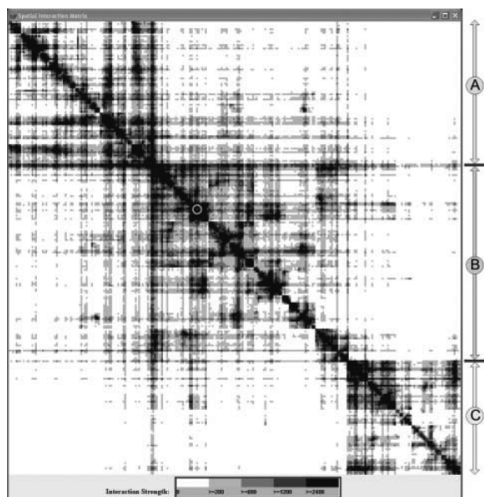
针对 OD 数据的地理空间点对结构、关系复杂、数量规模大、应用广泛等特点,本文分别从可视化方法、时空分析、视觉简化和领域应用 4 个方面对当前 OD 数据可视分析进行全面的调研与探讨.首先从 OD 数据的地理空间点对结构出发,阐述 OD 数据可视化原理和方法,介绍经典 OD 数据可视化方法分类及其适用场景;再从探索 OD 数据时空多维属性特征的角度出发,调研和分析 OD 数据特征的表达、感知、提取和呈现等方法,从空间分析和时空联合分析 2 个维度研究 OD 数据可视分析的效用和增强意义;进一步针对大规模 OD 数据可视化容易引起视觉混淆的问题,概括和介绍 OD 数据简化可视分析方法及其应用,包括过滤、采样和聚合等;然后调研和总结 OD 数据可视分析技术在不同行业领域的个性化应用方法及其定制的系统案例;最后,深入地讨论了 OD 数据可视分析研究中存在的挑战性问题和未来发展的趋势.

1 OD 数据可视化方法

OD 数据以 OD 点对的形式存储,与空间地图结合紧密.OD 点对通常包含空间位置及其流向关系,大量 OD 点对的空间位置和时间记录合成地理网络关系数据.如何有效地呈现 OD 数据及其在地图上的分布关系,是 OD 数据可视化研究的基础.传统的 OD 数据可视化方法主要包括 OD 矩阵、OD 嵌套地图和 OD 流图.

1.1 OD 矩阵

OD 矩阵类似于图的矩阵表示,如图 1 所示^[9].矩阵行表示起点,列表示终点,矩阵元素表示从起点到终点的 OD 点对,可以用颜色编码 OD 点对数量.为了处理大量的数据,重新排序和聚合技术^[8]也被应用于 OD 矩阵,以提高信息表达和可视化效率.

图 1 OD 矩阵^[9]

Ghoniem 等^[10]的一项用户研究发现, 邻接矩阵在快速解析邻接数据方面比节点链接图表现更好, 更适于对 OD 数据进行可视化. Dewulf 等^[11]利用 OD 矩阵呈现高峰和非高峰时间段公交车与私家车出行时间的区别. OD 矩阵对于区域之间的聚类关系表达比较清晰, 但在空间信息的表现上不够直观, 常常需要和地图结合起来以发现 OD 数据的地理位置及其空间分布规律. 例如, Ibarra 等^[12]提出连接条形图和协调图, 以支持 OD 矩阵的可视化分析, 同时保持空间位置关系.

1.2 OD 嵌套地图

OD 嵌套地图是一种网格化的二阶空间树形图, 利用嵌套的思想解决 OD 矩阵空间信息不直观的问题. 它将整个地图映射成 M 行 N 列的方格, 并在每个方格中嵌套一个相同的 M 行 N 列的地图, 每个 OD 点对可以映射为二阶小网格, 并用颜色编码流量, 如图 2 所示. 为了便于阅读, 网格的一些单元格可以留空, 以示意地图的轮廓形状.

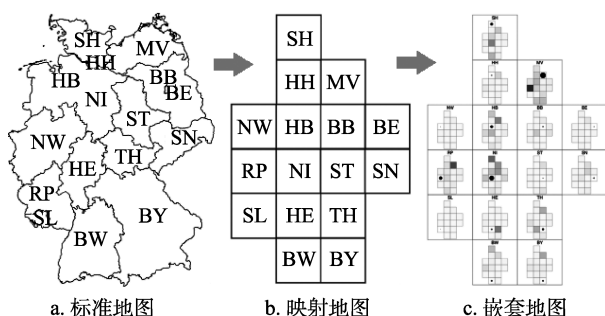
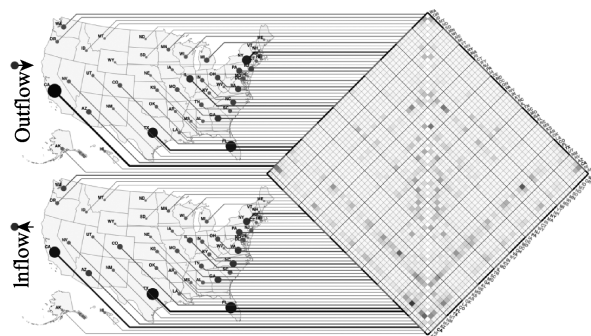


图 2 OD 嵌套地图

Wood 等^[13]通过构造网格化的 2 层空间树保留 OD 数据的空间布局, 利用 Hash 网格空间数据结构存储投影结果, 直观展现出 OD 数据的聚集情况

或流密度估计. 在此基础上, Yang 等^[14]设计了如图 3 所示 MapTriX, 通过矩阵旋转和连线重排算法, 叠加连线之间的约束关系, 以更清晰的方式向用户呈现了起终点间的流量关系.

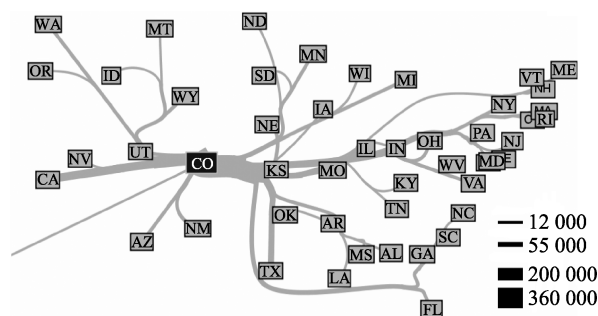
图 3 MapTriX^[14]

1.3 OD 流图

OD 流图是当前广泛使用的 OD 数据可视化手段, 即在地图上用线条连接起点和终点, 且线条的视觉元素可用于编码流量信息或其他属性信息^[15].

OD 流图源于 Tobler^[16]通过箭头表示美国 20 世纪 80 年代的人口迁移趋势地图, 在不丢失空间分辨率的情况下, 能够形象地解释迁移数据. Phan 等^[17]设计能自动生成流图的算法, 使流图成为 OD 数据可视化的主流方法, 主要可以分为一对多流图以及多对多流图 2 类.

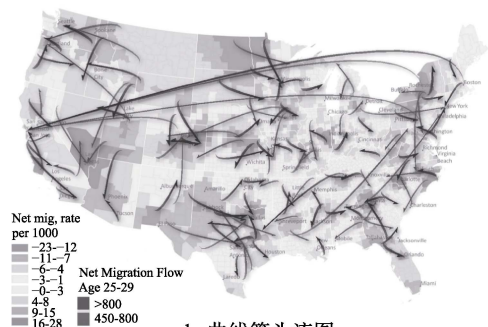
一对多流图即从起点开始流向多个终点的数, 通常采用树枝状流图的可视化形式. 图 4 所示为 1995—2000 年从美国科罗拉多州流出至各州的移民地图, 采用的是典型的树枝状流图形式^[17].

图 4 树枝状流图^[17]

多对多流图常用节点位置固定的网络可视化形式表达, 如图 5a 所示不同粗细的直线箭头^[7]和图 5b 所示不同颜色的曲线箭头^[18]. 另有学者专门研究箭头及曲线表达形式, 便于更加直观地理解流图^[19]. 在二维流图的基础上, Yang 等^[20]研究并设计 MapsLink 可视化图形, 更加快速和准确地实



a. 直线箭头流图



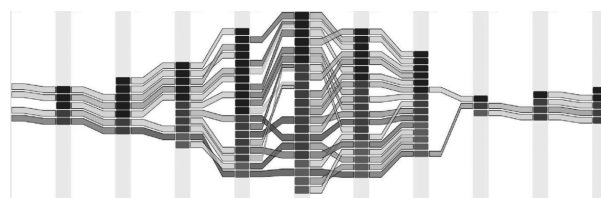
b. 曲线箭头流图

图 5 多对多流图^[18]

现三维地球和流图的协同可视化, 有效地解决复杂流图场景中的视觉混淆问题。

除以上 3 种常见 OD 数据可视化方法之外, Boyandin 等^[21]提出一种交互式的 OD 数据可视化方法 Flowstrates, 能够将 OD 数据起源和目的地显示在 2 个单独的映射中, 并且使用热力图表示流量大小随时间的变化. Guo 等^[22]设计开发 VIS-STAMP 系统, 整合计算、视觉和制图方法, 支持对复杂 OD 关联模式的概览, 并且通过各种交互方便用户关注和探索分析特定的模式. Lu 等^[23]提出的 OD 数据可视化工具 OD-Wheel, 帮助用户探索和挖掘 OD 数据, 识别和比较不同的 OD 数据模式. Wu 等^[24]提出移动转换图(mobility transition graph, MTG), 帮助用户探索和分析人群的时空转移特征, 如图 6 所示. 将移动模式表示为灰色矩形, 用颜色深浅编码模式出现的间隔次数, 相邻时间的移动模式转移用纽带连接表示, 纽带颜色用以编码模式间转移的概率。

综上所述, OD 矩阵图的表示方式丢失了地理位置信息, 视觉表达不直观; OD 嵌套地图尽可能地保留地理位置属性信息, 但是可视化结果不易解读分析; 相比 OD 矩阵和 OD 嵌套地图, OD 流图能够清晰直观地查看数据流向及规律. 然而, 考虑 OD 数据具有多维属性, 节点和边的关系复杂等特点, 简单的 OD 数据可视化方法难以发现数据内涵的复杂特征, 需要进一步从数据时空特征探索的角度提取和分析数据的分布规律和模式。

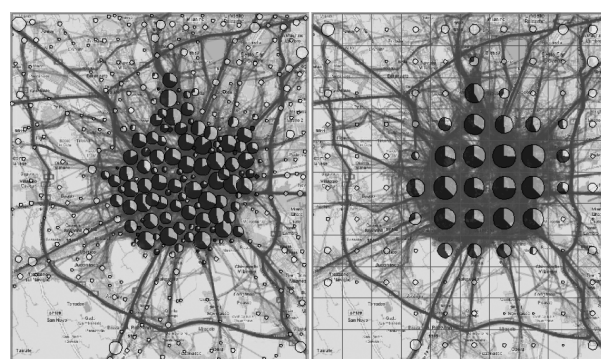
图 6 MTG^[24]

2 OD 数据时空特征可视分析

OD 数据在时间和空间维度上蕴含丰富的信息, 对其进行可视分析研究可以帮助用户从中发现特定的规律和模式. 本节从空间聚类、空间语义和时空联合 3 个角度对 OD 数据时空特征可视分析进行综述。

2.1 空间聚类分析

空间聚类分析是指将 OD 数据在空间维度上进行聚类可视分析. Zhu 等^[25]提出一种 OD 数据流聚类方法, 考虑起讫点以确定流的相似度, 最大限度地减少聚类过程中的信息损失, 并有效地改善流图遮挡和分布不均的视觉显示问题. Ferreira 等^[26]对连续 OD 轨迹进行聚类, 用基于向量场拟合的 k -means 算法, 将 OD 数据拟合至向量场模型, 能够有效地发掘数据隐含的运动模型. Verbeek 等^[27]提出利用边缘绑定计算高视觉质量的无交叉流, 基于使用对数螺旋的新型斯坦纳树诱导目标聚类和平滑束线. Rinzivillo 等^[28]设计一种空间渐进聚类技术, 可以在聚类的每步利用不同的距离函数分析处理具有异构属性的对象, 如图 7 所示。

图 7 空间渐进聚类^[28]

2.2 空间语义分析

空间语义分析是指对 OD 数据进行空间语义特征提取和描述, 进而获取并表征隐含的语义信息. Chu 等^[29]将地理坐标转换为反映上下文语义信息的街道名称, 通过文本主题建模技术识别隐藏

的出租车主题,进而有效地反映城市交通模式和趋势. Andrienko 等^[30]使用有向加权图表示运动数据,通过状态转换图表示语义高度抽象的运动行为视图,结合 OD 轨迹形式将运动的空间、时间和语义分析有机联结. Li 等^[31]通过 OD 数据与城市大数据(如社交媒体、交通运输、能源消耗等数据)相结合,提出一种时空语义立方体,旨在了解文本语义随空间和时间的变化. Wang 等^[32]提出一种利用空间语义识别功能区域的方法,使用兴趣点(point-of-interest, POI)数据和出租车轨迹数据研究验证了所提出的框架识别城市功能区的可行性,如图 8 所示. 该方法定量地探索功能区之间的相互作用,帮助用户更好地理解城市结构特征.

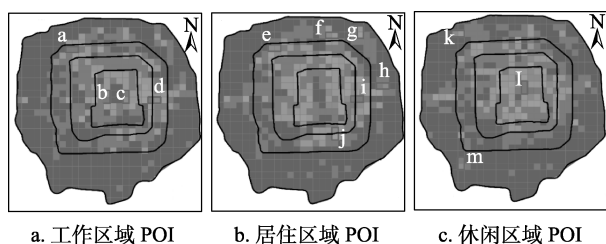


图 8 基于空间语义 POI 密度的功能区空间分布^[32]

2.3 时空联合分析

时空联合分析是指将数据的时间属性和空间属性联合起来进行分析,以探索数据的内在规律模式. Bai 等^[33]提出 FeatureFlow 可视化方法,将复杂的四维时空特征演化直观地显示在二维平面上,揭示包含动态和复杂特征的时空数据特征演变过程. Buchmuller 等^[34]提出一种空间高效的群组可视化技术 MotionRugs,在已建立空间划分策略的基础上,将每个时间步的空间维度缩减为单个实体的一维有序表示,进而有效地预测鱼群运动的时

序发展趋势. Kim 等^[35]利用重力模型提取时空变化,估计非定向时空数据在空间和时间上的连续分布,进而利用地图发现图 9 所示离散事件数据集的时空趋势,并利用全球定位系统(global positioning system, GPS)轨迹数据、推特数据、上海搜救事件和症状检测数据等数据集进行评估.

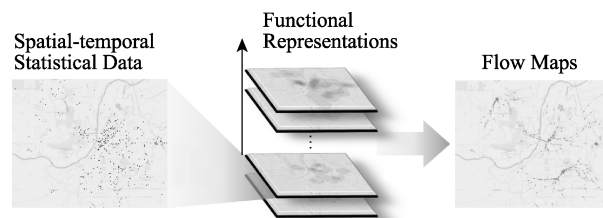


图 9 利用重力模型提取时间趋势流图^[35]

Li 等^[36]针对 OD 数据提出一个可视化的探索框架 COPE,有效地表达特定位置发生特定事件存在的特定时间关系,帮助用户探索和分析重要社会事件的形成和传播. Tominski 等^[37]提出混合的二维/三维显示可视化轨迹方法,在三维空间中叠加轨迹带,将具有时空特征的 OD 轨迹属性进行关联可视化,并在时空立方体中拓展了多维属性显示空间. Guo 等^[38]开发一个时空多维视角的可视分析系统,利用平行坐标显示多维交通轨迹数据,分别从空间、时间和多维视图角度交互检查数据,旨在关注微观交通模式与异常行为. Huang 等^[39]利用出租车轨迹数据和图模型创建 TrajGraph 可视分析系统,用直观的地图边点图和时间变化折线图呈现各个区域的道路使用情况,叠加热图在摘要中向用户显示隐藏的详细信息,支持用户以可视方式指定查询,帮助用户快速分析城市道路的重要程度和交通状况,如图 10 所示.

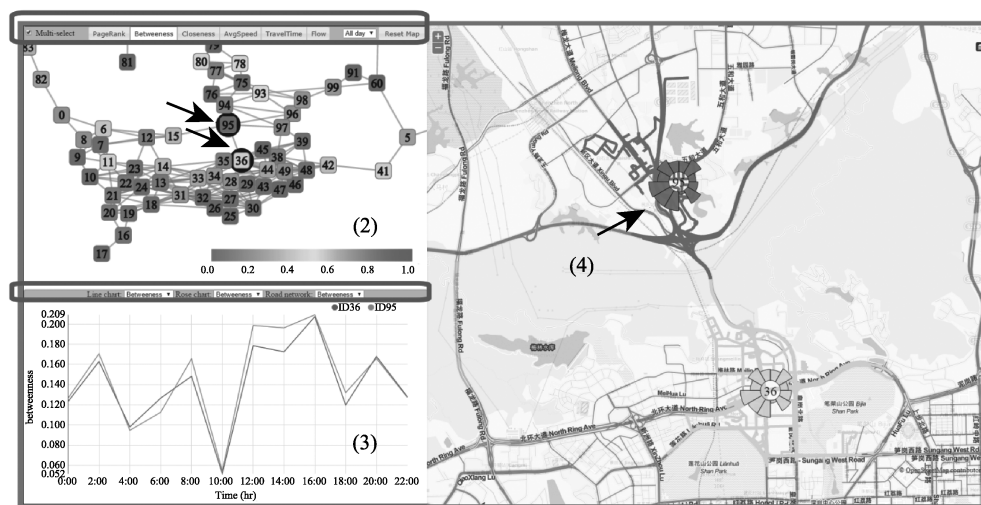


图 10 TrajGraph 系统界面^[39]

综上, 时空特征分析有利于 OD 数据的解析和探索, 有助于用户从数据中发现特定的规律和模式. 但随着数据规模的增加, 大规模 OD 数据的分析和展示面临巨大挑战, 传统的 OD 数据可视化方法往往因为流图中节点连线之间的重叠、交叉、覆盖等问题凸显, 严重影响视觉感知; OD 矩阵和 OD 嵌套地图也局限于展示篇幅和显示粒度, 难以从中探索内在联系和规律模式.

3 大规模 OD 数据视觉简化

简化是分析和呈现大规模 OD 数据的有效手段, 主要包括过滤、聚合和采样 3 类方法.

3.1 过滤

过滤通常是基于属性特征, 剔除或暂时隐藏部分数据, 将目标数据简化至用户感兴趣的范围, 以达到加快计算速度、提高分析效率等目标. 陆旻等^[40]针对地理空间数据所特有的空间信息, 概述

地理空间数据可视化中的过滤方法及相关工作. Zeng 等^[41]提出针对公共交通数据中途径限制的 OD 数据可视分析方法, 过滤筛选满足途径限制的轨迹, 设计能够有效探索 OD 模式的可视化方法. Wang 等^[42]针对稀疏的 OD 轨迹数据设计完成一个可视分析系统, 根据流量属性过滤 OD 轨迹数据, 重点关注宏观的交通模式. Krueger 等^[43]提出一种支持复杂过滤条件设置和长时间分析的交互式 OD 轨迹可视化工具, 使用透镜的可视化方式, 利用视觉的、基于集合的过滤器表达式选择 OD 轨迹. Hurter 等^[44]设计 FromDaDy 多维可视分析工具, 支持用户从视图中交互地过滤数据, 支持在目标视图上拖放以快速决定显示数据是否感兴趣, 进而增量地构建数据子集让用户实时评估所选内容的呈现效果. Lu 等^[45]面向出租车 OD 轨迹数据研制可视分析系统, 能够有效地分析和探索感兴趣区域的 OD 模式, 设计过滤模型允许用户迭代设置约束条件, 从而简化 OD 数据, 如图 11 所示.

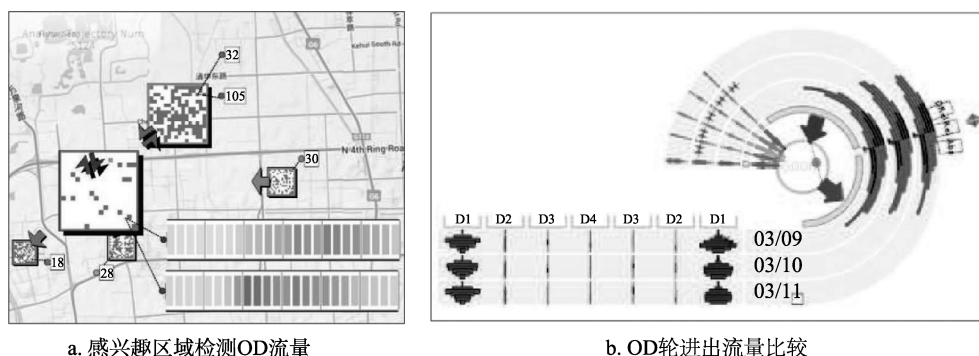


图 11 OD 数据点过滤^[45]

Harrower 等^[46]研制 ESV(earth systems visualizer)可视分析系统, 用户可以调整开始和结束日期对数据进行过滤, 增强对应的空间感知能力, 发现其内在关联, 如云层和空气温度之间的关系. Schmidt 等^[47]从时间角度对 OD 数据进行可视化呈现和分析, 允许用户从不同角度查看数据, 研究和测量随时间变化的空间变化, 达到向用户传达多维度信息的目的. van den Elzen 等^[48]提供交互式过滤和聚合的可视化工具, 如图 12 所示. 通过创建感兴趣的选择项, 并将网络中选择项的可视化与高层次的信息图概述相结合, 根据用户对原始网络图的操作生成抽象聚类结构, 将多维数据可视化方法嵌入聚类内部, 以展示选择项的结构和多变量属性.

3.2 采样

采样是计算机图形学中广泛使用的视觉简化

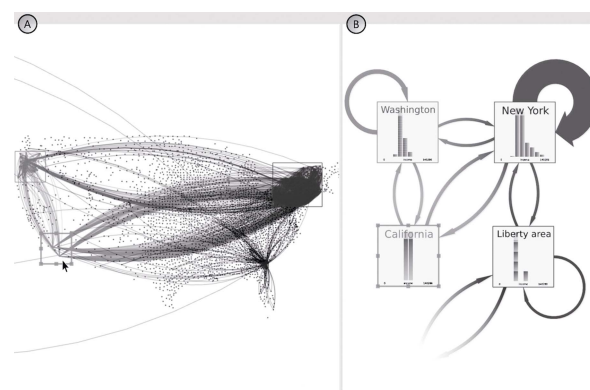


图 12 交互式过滤聚合系统^[48]

方法, 最常见的方法是随机采样. 对于多种数据维度的 OD 数据, 如何有效实现不同目标的特征保持, 是研究人员在选择采样策略时重点关注的目标之一. Zhou 等^[49]为减少 OD 流图的视觉混乱, 增加 OD 流之间的相关性, 保留更多语义信息, 设计

一个 OD 数据简化表达与可视分析系统, 在矢量化表征空间中设计迭代多目标的采样方案, 实现了 OD 特征视觉简化和语义特征保留采样. Chen 等^[50]对稀疏采样的带有地理标签的社交媒体数据进行

可视分析与探索, 利用不确定性模型通过错误数据过滤以及稀疏采样等技术得到有效的轨迹数据, 搭建一个分析社交媒体数据的交互式可视化系统, 实现空间结构保留采样, 如图 13 所示.

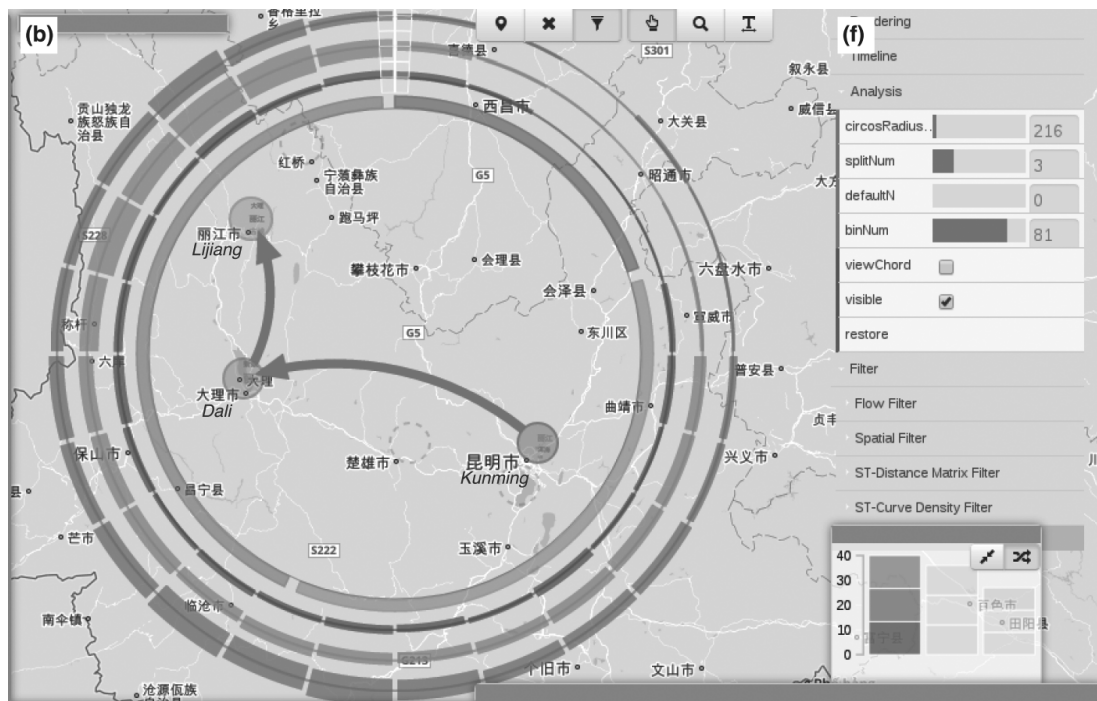


图 13 稀疏采样地理标签可视化系统^[50]

3.3 聚合

OD 数据聚合主要是指地理空间视觉元素的聚合, 有助于增强视觉感知, 发现地理模式和数据网络结构之间的关系, 是一种有效探索 OD 数据特征结构和多元关系的简化方法. 主要的聚合方法包括点聚合、边聚合和属性聚合 3 类.

点聚合^[51]是将空间毗连的起点或终点进行聚合, 以减少节点和节点间连线的复杂度, 呈现更清晰的可视分析结果. 图 14 所示为空间二维散点图的聚合示例. 金澄等^[52]依据轨迹点描述信息提取 OD 数据特征点, 利用分布密度指标和行政区划数据实施聚合分析, 对 OD 数据分布空间进行区域划分, 在揭示车流人流时空交互模式及辅助决策方面具有一定的贡献. Andrienko 等^[53]提出一种运动

数据的空间综合与聚合方法, 提取出轨迹的关键点进行空间聚合, 通过调整参数的方法控制聚合结果, 有效地应用于 OD 数据的交互式探索.

边聚合^[54]也称为边绑定, 是指将空间分布相近的数据流进行绑定聚合. Cui 等^[55]提出一种基于几何特征的边聚合框架, 将 OD 线条分组成束, 以减少整体的边交叉, 进而通过控制点约束 OD 网格的边聚合过程, 而且支持用户对边聚合结果进行交互. Ersoy 等^[56]提出了一种构造绑定布局的新方法, 利用二维骨架在形状中局部居中的特性, 使用位置信息相似的中间轴或骨骼作为绑定束的布局提示, 除了实现的快速性和简便性之外, 该方法还允许用户显式创建强分支或平滑绑定. Holten 等^[57]使用自组织方式进行边绑定, 即将边建模为可以相互吸引的柔性弹簧, 与以前的方法相比, 没有使用层次结构, 也没有控制网格, 所得的绑定视图能够有效降低视觉杂乱, 突出重要的边结构特征.

除了注重空间分布特征的点、边聚合方法, 另有系列聚合方法考虑属性特征对大规模 OD 数据进行简化. Andrienko 等^[58]通过聚合相关的运动事件来寻找运动数据中感兴趣的地点, 使用时空聚合来研

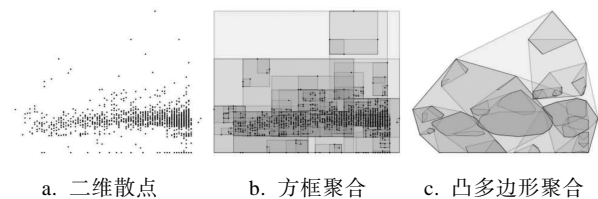


图 14 空间点的聚合

究兴趣点的移动数据的时间模式. Lampe 等^[59]发现对于动态流数据, 需要一个实时的核密度估计方法计算实现, 利用连续聚合的方法, 快速地更新核密度估计值, 增强数据的时间相关性.

聚合方法改变了原始数据的空间位置, 容易丢失局部细节的规律模式, 进而导致包括运行轨迹在内的信息丢失以及可更改区域单元问题(modifiable area unit problem, MAUP)^[60]. 针对这一缺点, Rae^[61]使用直线箭头测试了传统流图的可伸缩性限制, 试图通过覆盖的密度图使聚合流信息可见. Lyu 等^[62]提出一种新的交互式 OD 绑定技术, 在保持 OD 模式视觉简洁性的同时, 提高了对实际轨迹的地理空间信度. Graser 等^[63]探讨分析边缘绑定对空间数据集的适用性, 提出一种支持快速矩阵运算的边绑定方法, 并且基于地理信息系统对不同流动方向的边进行空间补偿绑定, 缩短边绑定的执行时间至传统方法的 60% 左右, 如图 15 所示.

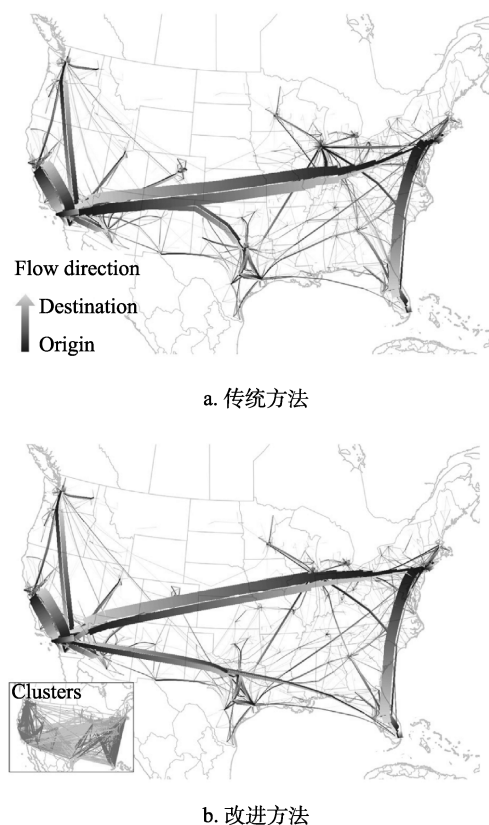


图 15 改进的边绑定方法^[63]

4 OD 数据可视化应用

社会生活各个领域都有 OD 数据产生, 对其进行可视分析有助于深入了解相关领域. 对于 OD 数

据, 可利用数学模型发现其中的隐含规律, 并应用到对应领域进行优化决策.

在智慧城市领域, Chen 等^[64]对交通数据可视化的基本概念、处理技术以及现有方法做了全面的综述, 根据交通数据的特点, 从时间、空间、时空和多变量 4 个方面总结了可视化的具体技术. Zeng 等^[65]根据大量乘客轨迹数据构建了公共交通系统移动模型, 从射频识别(radio frequency identification, RFID)卡的 OD 数据中推导估算出各种移动相关因素, 包括等待时间、骑行时间、传输时间和旅行效率等, 可视化分析乘客的移动规律. Liu 等^[66]将多维时空数据建模为张量, 利用算法切分数据并提取内在模式, 对区域销售趋势以及百货公司的客流量进行了可视分析和智能预测. Wang 等^[67]研制可视分析系统对城市交通拥堵状况进行视觉分析, 将 OD 数据清理后与路网匹配, 计算各路段的交通速度, 实现自动检测交通堵塞事件及其在时间和空间传播的描述. Liu 等^[68]根据出行时驾驶员可选择不同路线到达目的地的实际情形, 开发一种可视化技术以检测存在轨迹多样性的区域, 并显示与路径相关的高维属性和统计信息, 辅助管理者进行交通规划及决策. Liu 等^[69]从出租车轨迹数据中提取广告牌覆盖轨迹数、被覆盖轨迹路过广告牌的平均次数、总评分、平均费用等主要特征, 并在此基础上研制如图 16 所示广告牌放置点选择的可视分析系统 SmartAdP.

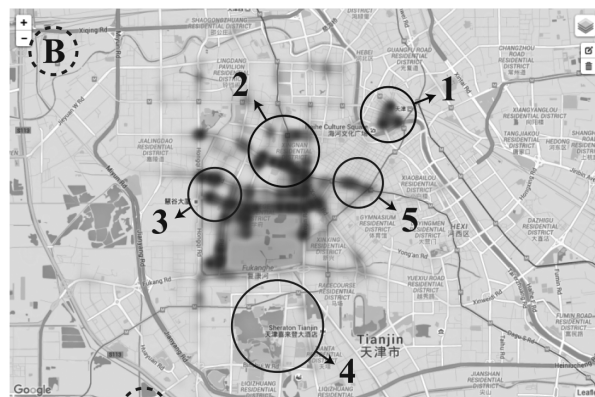


图 16 SmartAdP 系统^[69]

在社交媒体领域, Wu 等^[70]为了解城市人口流动中的共现现象, 分析了来自多个地区的人们在同一时间段访问不同地方行为中的相关性. Gonzalez 等^[71]利用 10 万多名匿名手机用户的 OD 数据, 发现人类轨迹具有极高的时空规律性, 除去旅行和外出等常规迁移, 人类依然遵循简单的可重复模式, 这一发现对于城市规划、交通预测和疾病防

控等领域具有重要意义. Chen 等^[72]提出一种具有地图隐喻的可视分析方法 R-Map, 支持对社交媒体中语义信息重新发布过程的调查和分析, 并提供交互用于对信息扩散过程中的时间模式和用户行为进行探索. Wang 等^[73]研制 IdeaFlow 视觉分析系统追踪相关思想在多个社会群体内部和之间的 OD 流轨迹, 并分析其超前滞后的变化效果, 探索不同时间粒度级别的大量数据.

在生态卫生领域, Ramirez-Hernandez 等^[74]采集新鲜的马匹粪便样本, 使用 OD 流图对马匹的运动进行可视化, 探究可能感染工作马匹的寄生虫物种的重要性和多样性及其对劳动者的健康和福祉的影响. Qi 等^[75]利用人口普查和能源统计数据计算并可视化我国人口迁移过程中的碳排放流量 OD 数据, 探索碳排放转移方向的规律, 为促进低碳行为和优化能源政策做出有益贡献. Konzack 等^[76]建立 75 只海鸥长达 3 年的迁徙 OD 数据集, 可识别中途停留地点、环境条件或具有特定迁移行为的个体或群体, 支持生态学家的分析工作.

5 结 语

OD 数据可视分析是数据可视化、地理空间分析等领域的研究热点, 也是社会各行业的重要应用技术. 本文阐述了经典的 OD 数据可视化方法, 并从时空特征表达、感知、提取和探索分析角度, 以及从空间维度和时空联合维度对 OD 数据进行可视分析; 进一步梳理了大规模 OD 数据可视分析研究方法中过滤、采样和聚合等视觉简化方法及其应用; 最后调研了 OD 数据可视分析技术在不同行业领域的应用及具体案例. 结合当前 OD 数据可视分析的进展, 本文进一步总结和展望 OD 数据可视分析研究中存在的挑战和未来发展的趋势.

(1) 随着数据采集过程的规范化和采集粒度的不断细化, 大规模 OD 数据快速地产生和累积. 如何有效地融合数据挖掘模型和可视化技术, 快速、准确地从大规模 OD 数据中提取和表示用户感兴趣的特征是大规模 OD 数据探索和分析的关键.

(2) 随着 OD 数据规模的增加, OD 数据可视化方法存在视觉冗余、拓展性不强等问题, 而聚类 and 简化等方法容易损失局部关键细节特征, 具有一定的局限性. 因此, 面向大规模 OD 数据, 开发细节特征保留的简化可视分析方法, 对于大规模 OD 数据的全局分析和细节探索具有重要意义.

(3) OD 数据是典型的稀疏数据, 对其进行特征表达和可视分析难免会存在信息损失和与实际不符合的情况. 因此, 如何利用可视化技术直观地引导更加准确的 OD 稀疏数据还原过程, 且在误差不可避免的前提下提供必要的不确定性分析和展示方法, 是面向 OD 数据开展探索和分析研究所面临的重要挑战.

参考文献(References):

- [1] Andrienko G, Andrienko N, Fuchs G, *et al.* Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 23(9): 2120-2136
- [2] Andrienko N, Andrienko G. Visual analytics of movement: an overview of methods, tools and procedures[J]. *Information Visualization*, 2013, 12(1): 3-24
- [3] Pu Jiansu, Qu Huamin, Ni Mingxuan. Survey on visualization of trajectory data[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2012, 24(10): 1273-1282(in Chinese)
(蒲剑苏, 屈华民, 倪明选. 移动轨迹数据的可视化[J]. *计算机辅助设计与图形学学报*, 2012, 24(10): 1273-1282)
- [4] Wang Zuchao, Yuan Xiaoru. Visual analysis of trajectory data[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2015, 27(1): 9-25(in Chinese)
(王祖超, 袁晓如. 轨迹数据可视分析研究[J]. *计算机辅助设计与图形学学报*, 2015, 27(1): 9-25)
- [5] Zhou Zhiguang, Shi Chen, Shi Linsong, *et al.* A survey on the visual analytics of geospatial data[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2018, 30(5): 747-763(in Chinese)
(周志光, 石晨, 史林松, 等. 地理空间数据可视分析综述[J]. *计算机辅助设计与图形学学报*, 2018, 30(5): 747-763)
- [6] Xie Chong, Guan Xuefeng, Zhou Weixuan, *et al.* MAP-Vis: a spatio-temporal big data visualization method based on multi-dimensional aggregation[J]. *Journal of Image and Graphics*, 2019, 24(5): 816-826(in Chinese)
(谢冲, 关雪峰, 周炜轩, 等. MAP-Vis: 基于 MAP 模型的时空点状大数据可视化方案[J]. *中国图象图形学报*, 2019, 24(5): 816-826)
- [7] Guo D S. Flow mapping and multivariate visualization of large spatial interaction data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2009, 15(6): 1041-1048
- [8] Guo D S, Gahegan M. Spatial ordering and encoding for geographic data mining and visualization[J]. *Journal of Intelligent Information Systems*, 2006, 27(3): 243-266
- [9] Guo D. Visual analytics of spatial interaction patterns for pandemic decision support[J]. *International Journal of Geographical Information Science*, 2007, 21(8): 859-877
- [10] Ghoniem M, Fekete J D, Castagliola P. A comparison of the

- readability of graphs using node-link and matrix-based representations[C] //Proceedings of the IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2004: 17-24
- [11] Dewulf B, Neutens T, Vanlommel M, *et al.* Examining commuting patterns using floating car data and circular statistics: exploring the use of new methods and visualizations to study travel times[J]. *Journal of Transport Geography*, 2015, 48: 41-51
- [12] Ibarra J C, Triana J A, Hernandez J T. Visualization of origin-destination matrices using a connection barchart and coordinated maps[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2016: 189-193
- [13] Wood J, Dykes J, Slingsby A. Visualisation of origins, destinations and flows with OD maps[J]. *The Cartographic Journal*, 2010, 47(2): 117-129
- [14] Yang Y L, Dwyer T, Goodwin S, *et al.* Many-to-many geographically-embedded flow visualisation: an evaluation[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 23(1): 411-420
- [15] Koylu C, Guo D S. Design and evaluation of line symbolizations for origin-destination flow maps[J]. *Information Visualization*, 2016, 16(4): 309-331
- [16] Tobler W R. Experiments in migration mapping by computer[J]. *The American Cartographer*, 1987, 14(2): 155-163
- [17] Phan D, Xiao L, Yeh R, *et al.* Flow map layout[C] // Proceedings of the IEEE Symposium on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2005: 219-224
- [18] Guo D S, Zhu X. Origin-destination flow data smoothing and mapping[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2014, 20(12): 2043-2052
- [19] Jenny B, Stephen D M, Muehlenhaus I, *et al.* Design principles for origin-destination flow maps[J]. *Cartography and Geographic Information Science*, 2016, 45(1): 62-75
- [20] Yang Y L, Dwyer T, Jenny B, *et al.* Origin-destination flow maps in immersive environments[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(1): 693-703
- [21] Boyandin I, Bertini E, Bak P, *et al.* Flowstrates: an approach for visual exploration of temporal origin-destination data[J]. *Computer Graphics Forum*, 2011, 30(3): 971-980
- [22] Guo D S, Chen J, MacEachren A M, *et al.* A visualization system for space-time and multivariate patterns(VIS-STAMP)[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(6): 1461-1474
- [23] Lu M, Wang Z C, Liang J, *et al.* OD-Wheel: visual design to explore OD patterns of a central region[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2015: 87-91
- [24] Wu F R, Zhu M F, Wang Q, *et al.* Spatial-temporal visualization of city-wide crowd movement[J]. *Journal of Visualization*, 2017, 20(2): 183-194
- [25] Zhu X, Guo D S. Mapping large spatial flow data with hierarchical clustering[J]. *Transactions in GIS*, 2014, 18(3): 421-435
- [26] Ferreira N, Klosowski J T, Scheidegger C E, *et al.* Vector field *k*-means: clustering trajectories by fitting multiple vector fields[J]. *Computer Graphics Forum*, 2013, 32(3): 201-210
- [27] Verbeek K, Buchin K, Speckmann B. Flow map layout via spiral trees[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2011, 17(12): 2536-2544
- [28] Rinzivillo S, Pedreschi D, Nanni M, *et al.* Visually driven analysis of movement data by progressive clustering[J]. *Information Visualization*, 2008, 7(3): 225-239
- [29] Chu D, Sheets D A, Zhao Y, *et al.* Visualizing hidden themes of taxi movement with semantic transformation[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2014: 137-144
- [30] Andrienko N, Andrienko G. State transition graphs for semantic analysis of movement behaviours[J]. *Information Visualization*, 2018, 17(1): 41-65
- [31] Li J, Chen S M, Chen W, *et al.* Semantics-space-time cube: a conceptual framework for systematic analysis of texts in space and time[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2020, 26(4): 1789-1806
- [32] Wang Y D, Gu Y Y, Dou M X, *et al.* Using spatial semantics and interactions to identify urban functional regions[J]. *ISPRS International Journal of Geo-Information*, 2018, 7(4): 130-145
- [33] Bai Z H, Tao Y B, Lin H. FeatureFlow: exploring feature evolution for time-varying volume data[J]. *Journal of Visualization*, 2019, 22(5): 927-940
- [34] Buchmuller J, Jackie D, Cakmak E, *et al.* MotionRugs: visualizing collective trends in space and time[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(1): 76-86
- [35] Kim S, Jeong S, Woo I, *et al.* Data flow analysis and visualization for spatiotemporal statistical data without trajectory information[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 24(3): 1287-1300
- [36] Li J, Chen S M, Zhang K, *et al.* COPE: interactive exploration of co-occurrence patterns in spatial time series[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(8): 2554-2567
- [37] Tominski C, Schumann H, Andrienko G, *et al.* Stacking-based visualization of trajectory attribute data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2565-2574
- [38] Guo H Q, Wang Z C, Yu B W, *et al.* Tripvista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2011: 163-170
- [39] Huang X K, Zhao Y, Yang J, *et al.* TrajGraph: a graph-based visual analytics approach to studying urban network centralities using taxi trajectory data[J]. *IEEE Transactions on*

- Visualization and Computer Graphics, 2015, 22(1): 160-169
- [40] Lu Min, Yuan Xiaoru. Filter in visualization of geospatial data[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(5): 702-711(in Chinese)
(陆旻, 袁晓如. 地理空间数据可视化中的过滤[J]. 计算机辅助设计与图形学学报, 2016, 28(5): 702-711)
- [41] Zeng W, Fu C W, Müller Arisona S, *et al.* Visualizing waypoints-constrained origin destination patterns for massive transportation data[J]. Computer Graphics Forum, 2016, 35(8): 95-107
- [42] Wang Z C, Ye T Z, Lu M, *et al.* Visual exploration of sparse traffic trajectory data[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 1813-1822
- [43] Krueger R, Thom D, Woerner M, *et al.* TrajectoryLenses - a set-based filtering and exploration technique for long-term trajectory data[J]. Computer Graphics Forum, 2013, 32(3): 451-460
- [44] Hurter C, Tissoires B, Conversy S. FromDaDy: spreading aircraft trajectories across views to support iterative queries [J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15(6): 1017-1024
- [45] Lu M, Liang J, Wang Z C, *et al.* Exploring OD patterns of interested region based on taxi trajectories[J]. Journal of Visualization, 2016, 19(4): 811-821
- [46] Harrower M, MacEachren A, Griffin A L. Developing a geographic visualization tool to support earth science learning [J]. Cartography and Geographic Information Science, 2000, 27(4): 279-293
- [47] Schmidt J, Fleischmann D, Preim B, *et al.* Popup-plots: warping temporal data visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(7): 2443-2457
- [48] van den Elzen S, van Wijk J J. Multivariate network exploration and presentation: from detail to overview via selections and aggregations[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 2310-2319
- [49] Zhou Z G, Meng L H, Tang C, *et al.* Visual abstraction of large scale geospatial origin-destination movement data[J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 25(1): 43-53
- [50] Chen S M, Yuan X R, Wang Z H, *et al.* Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1): 270-279
- [51] Elmqvist N, Fekete J D. Hierarchical aggregation for information visualization: overview, techniques, and design guidelines[J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 16(3): 439-454
- [52] Jin Cheng, Chen Yuanyuan, Yang Min. A multi-scale visualization method for the trajectory origin-destination data[J]. Journal of Geo-Information Science, 2017, 19(8): 1011-1018(in Chinese)
(金澄, 陈媛媛, 杨敏. 面向轨迹起止特征点数据的多比例尺可视化方法[J]. 地球信息科学学报, 2017, 19(8): 1011-1018)
- [53] Andrienko N, Andrienko G. Spatial generalization and aggregation of massive movement data[J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(2): 205-219
- [54] Zeng W, Shen Q, Jiang Y, *et al.* Route-aware edge bundling for visualizing origin-destination trails in urban traffic[J]. Computer Graphics Forum, 2019, 38(3): 581-593
- [55] Cui W, Zhou H, Qu H, *et al.* Geometry-based edge clustering for graph visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2008, 14(6): 1277-1284
- [56] Ersoy O, Hurter C, Paulovich F, *et al.* Skeleton-based edge bundling for graph visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 17(12): 2364-2373
- [57] Holten D, van Wijk J J. Force-directed edge bundling for graph visualization[J]. Computer Graphics Forum, 2009, 28(3): 983-990
- [58] Andrienko G, Andrienko N, Hurter C, *et al.* Scalable analysis of movement data for extracting and exploring significant places[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(7): 1078-1094
- [59] Lampe O D, Hauser H. Interactive visualization of streaming data with kernel density estimation[C] //Proceedings of the IEEE Pacific Visualization Symposium. Los Alamitos: IEEE Computer Society Press, 2011: 171-178
- [60] von Landesberger T, Brodtkorb F, Roskosch P, *et al.* Mobility graphs: visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1): 11-20
- [61] Rae A. From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census[J]. Computers, Environment and Urban Systems, 2009, 33(3): 161-178
- [62] Lyu Y, Liu X, Chen H Y, *et al.* OD morphing: balancing simplicity with faithfulness for OD bundling[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(1): 811-821
- [63] Graser A, Schmidt J, Roth F, *et al.* Untangling origin-destination flows in geographic information systems[J]. Information Visualization, 2019, 18(1): 153-172
- [64] Chen W, Guo F Z, Wang F Y. A survey of traffic data visualization[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(6): 2970-2984
- [65] Zeng W, Fu C W, Arisona S M, *et al.* Visualizing mobility of public transportation system[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 1833-1842
- [66] Liu D Y, Xu P P, Ren L. TPFlow: progressive partition and multidimensional pattern extraction for large-scale spatio-temporal data analysis[J]. IEEE Transactions on Visualization and Computer Graphics, 2019, 25(1): 1-11
- [67] Wang Z C, Lu M, Yuan X R, *et al.* Visual traffic jam analysis based on trajectory data[J]. IEEE Transactions on Visualization

- and Computer Graphics, 2013, 19(12): 2159-2168
- [68] Liu H, Gao Y, Lu L, *et al.* Visual analysis of route diversity[C] //Proceedings of the IEEE Conference on Visual Analytics Science and Technology. Los Alamitos: IEEE Computer Society Press, 2011: 171-180
- [69] Liu D Y, Weng D, Li Y H, *et al.* SmartAdP: visual analytics of large-scale taxi trajectories for selecting billboard locations[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1): 1-10
- [70] Wu W C, Xu J Y, Zeng H P, *et al.* TelCoVis: visual exploration of co-occurrence in urban human mobility based on telco data[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1): 935-944
- [71] Gonzalez M C, Hidalgo C A, Barabasi A L. Understanding individual human mobility patterns[J]. Nature, 2008, 453 (7196): 779-782
- [72] Chen S, Li S H, Chen S M, *et al.* R-Map: a map metaphor for visualizing information reposting process in social media[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 26(1): 1204-1214
- [73] Wang X T, Liu S X, Chen Y, *et al.* How ideas flow across multiple social groups[C] //Proceedings of the IEEE Conference on Visual Analytics Science and Technology. Los Alamitos: IEEE Computer Society Press, 2016: 51-60
- [74] Ramirez-Hernandez A, Polo G, Robayo-Sanchez L N, *et al.* Gastrointestinal and pulmonary parasites of working horses from Colombia[J]. Veterinary Parasitology, Regional Studies and Reports, 2019, 17: 100296
- [75] Qi W, Li G D. Residential carbon emission embedded in China's inter-provincial population migration[J]. Energy Policy, 2020, 136: 111065
- [76] Konzack M, Gijsbers P, Timmers F, *et al.* Visual exploration of migration patterns in gull data[J]. Information Visualization, 2019, 18(1): 138-152