引文格式：揭滢萤,陆妍玲,施钰琪,等. 融合轨迹特征的多要素时空大数据组织方法[J]. 地理空间信息, 2026, 24(3): 7-11. doi:10. 3969/j.issn. 1672-4623.2026.03.002

融合轨迹特征的多要素时空大数据组织方法

揭滢萤[1]，陆妍玲[1]，施钰琪[1]，傅　毓[2*]，陆深传[2]
（1. 桂林理工大学 测绘地理信息学院，广西 桂林 541004；2. 广西壮族自治区地理信息测绘院，广西 柳州 545005）

摘　要：针对传统大数据组织方法难以满足多模态、多尺度、多维度时空大数据管理的难题，提出了一种融合轨迹特征的多要素时空大数据组织方法。融合时间、空间、属性、轨迹和关系等多个要素对时空大数据进行建模，重点剖析了不同要素的可视化表达及其关联。基于该模型，对收集的182个用户6年17 621条GPS轨迹案例数据进行可视化分析，提取其轨迹特征；并结合DBSCAN聚类方法进行应用。结果表明，该模型能有效融合轨迹特征，对时空大数据进行有效存储管理，并实现更深入的数据挖掘。

关键词：时空大数据；轨迹特征；多要素模型；聚类分析
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2026）03-0007-05

Multi-element Spatio-temporal Big Data Organization Method by Integrating Trajectory Features

JIE Yingying[1], LU Yanling[1], SHI Yuqi[1], FU Yu[2], LU Shenchuan[2]
(1. College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China;
2. Guangxi Geographic Information Mapping Institute, Liuzhou 545005, China)

Abstract: To address the challenges of traditional big data organization methods that primarily focused on single data management tasks and struggled with managing multi-modal, multi-scale, and multi-dimensional spatio-temporal big data, we proposed a multi-element spatio-temporal big data organization method integrating trajectory features. We constructed spatio-temporal big data model by integrating multi-element such as time, space, attributes, trajectories and relationships, and analyzed the visual representation of different elements and their interconnections. Based on this model, we conducted visual analysis on collected case data comprising 17 621 GPS trajectories from 182 users over 6 years, extracted trajectory features, and applied it in combination with the DBSCAN clustering method. The result shows that the proposed model effectively integrates trajectory features, enables efficient storage management of spatio-temporal big data, and facilitates deeper data mining.

Key words: spatio-temporal big data, trajectory feature, multi-element model, clustering analysis

时空大数据技术的快速发展给各行业的生产研究带来了极大的便利。据中国产业信息网分析，未来我国大数据市场需求将保持35%的速度增长[1-2]。要高效应用时空大数据，就必须对其进行科学的组织管理。目前数据超体量、结构复杂是时空大数据组织与管理的技术瓶颈[3-4]。随着我国科技实力的稳步发展以及综合国力的不断增强，对时空大数据的应用需求越来越高，有效管理组织时空大数据成为当前的迫切需求[5]。部分专家学者[6-8]对时空大数据平台提出了各自的见解；而轨迹作为时空大数据的鲜明特征，也吸引了部分学者[9]研究相关内容。时空大数据的多模态、多尺度特性对传统数据组织方法提出了严峻挑战，尤其是在多要素协同建模与动态轨迹融合方面亟待突破。具体表现为：①现有模型难以统一表达时间、空间、属性、轨迹、关系等多要素关联；②动态轨迹数据的异频采样与多源异构特性导致要素间时空语义关联挖掘不足；③缺乏高效的聚类与可视化方法支撑多要素数据的实时分析与应用。鉴于此，本文提出了一种融合轨迹特征的多要素时空大数据组织方法，实现了动态数据的结构化封装与关联表达；结合DBSCAN密度聚类算法优化轨迹交叉区域的要素关系挖掘；基于182个用户长达6年的GPS轨迹数据，验证了模型在数据组织与深度挖掘中的有效性，为社会管理、智慧城市、公共安全等场景提供技术支撑。

项目来源：国家自然科学基金资助项目（41961063）。
第一作者简介：揭滢萤（2003—），本科，研究方向为GIS数据分析，E-mail：3212080701211@glut.edu.cn。
通信作者：傅毓（1986—），硕士，高级工程师，从事GIS工程开发工作，E-mail：fuy2010@126.com。

1 时空大数据多要素定义

时空大数据数据量大、种类多、增长快，既具备地理空间的静态特征（如时间、空间和属性），又具备由动态性衍生出轨迹、动态关系等独特特征，这些特征共同支撑了时空大数据的多要素定义。

1.1 时间要素

地理实体是地理空间的主要描述对象，同一地理实体或地理现象的数据信息将随时间的变化而变化[10]。将记录和描述地理事件发生时间的最小单元称为时间要素单元，一个地理事件的发生可由多个时间要素单元堆叠表达。相较于空间的三维性，时间具有一维性、流动性和不可重复性。时间要素的数学表达为：

$$\Delta t \in T=[t_0, t] \tag{1}$$

$$D_t=\begin{cases} h(t)=t_0+n\Delta t \\ d_t=d_0+d_{h(t)} \end{cases} \tag{2}$$

式中，Δt 为时间要素单元；$[t_0, t]$ 为时间集合；D_t 为多个 Δt 堆叠区间产生的时空数据；$h(t)$ 为时间函数；t_0 为数据开始时刻；$n\Delta t$ 为开始至结束的变化区间；d_t 为当前时间下包含的时空数据；d_0 为 t_0 时刻数据状态；$d_{h(t)}$ 为在时间要素影响下发生改变的时空数据。

在一维轴上时间要素单元表示的是某个时间点；在二维平面上时间要素需要与其他要素耦合，才能准确表征时空数据的时序变化特征；在三维空间上时间要素则需与更多要素进行耦合来描述变化。

1.2 空间要素

现实地理空间本质上是一个三维空间，各种运动中地理实体或地理现象的空间位置会发生改变，进而导致其时空大数据发生变化。因此，通过空间要素描述地理实体在空间上的数据详细程度。时空大数据的空间要素变化主要表现在空间位置以及空间对象间空间关系的变化。空间要素的定义为：

$$S=\begin{cases} S_l=S_l(X, Y, Z) \\ S_r=sr(a)+sr(b)+\cdots+sr(i) \end{cases} \tag{3}$$

式中，S_l 为空间位置要素；S_r 为空间关系要素；$sr(a)$ 为两个及以上相关地理实体包含的空间关系。

1.3 属性要素

属性是对地理实体性质的描述，是客观存在的现象，不会自动抽象表现为可视化信息。时空大数据中的属性数据记录了地理实体或现象的全部信息及其附属相关信息[11-12]，因此需针对地理实体或现象创建属性要素。基于地理实体的时空特征，可将属性要素分为静态属性和动态属性。由于动态属性将造成地理实体发生量变，累积则会导致质变。若对属性要素进行记录，则只需记录地理实体对象发生改变的动态属性，无需重复记录静态属性，可在一定程度上减少数据冗余，提高描述的精准度。时空大数据的属性信息组织需要信息准确、完整，并与地理实体或地理现象一一对应。

1.4 轨迹要素

时空大数据具有动态性，其动态演化的结果在很多方面都产生了对应的轨迹数据。现实地理实体的运动投影在地理空间中便产生了相应的轨迹数据，通常由地理空间坐标集和时间表示，如 $P_1 \to P_2 \to P_3 \to \cdots \to P_n$，其中 $P=(x, y, t)$，x、y 为大地坐标系中的经纬度，t 为时间。轨迹数据是指时空域中存在连续性的序列集[13]，如车辆移动信息、社交平台中用户签到信息等。基于时空大数据的轨迹要素具有以下特征：①时空序列性，在时间上具有连续性，空间上具有相似性，整体呈现序列性；②异频采样性，时空大数据的轨迹要素产生的个体不同，具有较大的随机性和不确定性，而各平台或机构采集数据的频率间隔也有差别；③多源异构性，数据来源、类型各不相同。根据时空大数据轨迹要素的特点，某一地理实体或现象的轨迹由一系列按照时间排序的轨迹点组成。轨迹要素描述了某一地理实体或现象在运动过程中产生的时空大数据，是研究移动对象最常见的数据结构[14]。

1.5 关系要素

地理实体对象在客观真实世界中是不断变化的，时空大数据不仅记录了地理实体个体随时间空间变化而发生的变化，还包含了实体间千丝万缕的联系。要完整描述地理实体对象的时空变化，则需将关系要素与时间要素、空间要素、属性要素和轨迹要素进行组合表达。由于时间要素具有一维流动性，因此空间要素和属性要素都是基于时间要素对地理实体的描述。一旦空间和属性要素发生改变，就可能带动一个或多个关系发生变化，因此关系要素在时空演变分析过程中具有不可或缺的作用。

2 融合轨迹特征的时空大数据多要素模型构建

2.1 时空大数据要素化

时空大数据来源于地理实体，在模型构建过程中，采用面向对象的方法对现实地理实体或地理现象进行抽象表达。将地理实体或现象抽象为单一实体的单一对象、两个或两个以上实体的复合对象以及任意实体组合的空间对象；再根据实际需求对对象进行分

类。根据多要素关系的定义，对地理实体的不同要素进行封装，每个要素类都封装了自己独立的属性和行为，能独立完成一些操作[15]。在实体的基础上对时空大数据的多要素进行封装，但多要素彼此间通过时间和空间关联，这样不仅使时空大数据更具针对性，而且在数据结构方面也增加了便利性。多要素各要素类也是由不同时段的数据组成：时间要素 C_{Ti} 由 $Ti_i(i=1,2,3,\cdots,n)$ 组成；空间要素 C_{Sp} 通常恒定，但有时根据需求变化，由 $Sp_i(i=1,2,3,\cdots,n)$ 组成；属性要素 C_{At} 由 $At_i(i=1,2,3,\cdots,n)$ 组成；轨迹要素 C_{Lo} 由 $Lo_i(i=1,2,3,\cdots,n)$ 组成；关系要素 C_{Op} 由 $Op_i(i=1,2,3,\cdots,n)$ 组成。由此，基于多要素特征的时空大数据模型可表示为：

$$\boldsymbol{TS}=\begin{bmatrix}C_{At}\\C_{Lo}\\C_{Op}\end{bmatrix}\begin{bmatrix}C_{Ti} & C_{Sp}\end{bmatrix}=\begin{bmatrix}C_{At}C_{Ti} & C_{At}C_{Sp}\\C_{Lo}C_{Ti} & C_{Lo}C_{Sp}\\C_{Op}C_{Ti} & C_{Op}C_{Sp}\end{bmatrix} \tag{4}$$

地理实体在时空变化过程中，轨迹和关系要素并不是独立的。多个地理实体的轨迹可能重叠或交叉，关系要素也是多个地理实体相互关联所产生的关系。在轨迹重叠或交叉处，利用聚类算法可对当前时空要素下产生关联的地理实体进行计算，并挖掘其关联关系要素。DBSCAN是一种典型的基于密度的聚类算法，能将足够高密度的区域划分为簇，并在具有噪声的空间数据库中发现任意形状的簇。因此可通过DBSCAN算法对时空大数据进行聚类计算。首先根据实际情况选定聚类中心 P_ε（通常为轨迹交叉处），设定邻域半径Eps(ε)，以 P_ε 为球心，Eps(ε)为半径的球体内包含的地理实体数量即为该簇的密度。邻域 ε 内至少包含MinPts个地理实体，即 $N_\varepsilon(P_\varepsilon)\geqslant$MinPts。簇内包含的地理实体即为标定时间点处可能互相产生交集关联的实体，结合轨迹要素和关系要素，有助于判定指定时间点通过指定空间点的地理时空实体数量及属性，如新冠疫情中可将某一感染者的活动数据要素化，以其活动轨迹为中心轨迹，可视化同一时间同一空间的其他人类活动轨迹，寻找交点，设置筛查邻域半径进行聚类，筛选出其密切接触者进一步分析。

2.2　融合轨迹特征的时空大数据多要素模型

针对面向地理实体的时空大数据，根据地理实体的组成特点，采集地理实体的各项数据；再按照多要素特征对数据进行分类封装，构成多要素模型。该模型有利于数据结构的完整性，在查询和利用方面具有较强的针对性，且对实体和多要素具有较强的抽象表达和完整的数据结构。创建时空大数据的多要素模型，通过模型链接实体对象与时空大数据，各要素信息具有各自的操作方法又相互关联，符合时空大数据的动态演化过程。多要素模型通过接口实现实体与外界的联系，使得时空大数据的多要素信息可根据不同场景发生变化。对实体对象和多要素进行抽象表达和封装，各要素包含了实体对象的相应要素信息和操作，具有不同的识别码，因此多要素模型采用BNF范式形式抽象化表达，即<Object, Ti, Sp, At, Lo, Op>。逻辑模型见图1，该框架表达了地理空间中时空大数据空间对象、几何元素以及多要素之间的逻辑关系。将实体对象虚拟化提取为多源异构的时空大数据对地理实体进行组织管理。将时空大数据进行要素化抽象和表示，根据各要素特点将信息归类为相应的要素类别。各要素与空间对象之间通过要素标识进行时空关联。地理实体的时间要素和空间要素都发生变化后，记录变化后的时间要素和空间要素，将其作为标识点；再记录标识点的属性要素和关系要素；然后将一系列标识点连接起来记录轨迹要素，以此表达地理实体的多要素模型时空演变过程。首先对地理实体进行虚拟数据化；再根据时空大数据的要素变化描述地理空间实体的变化；然后结合时空大数据的多要素时空关联分析，建立时空大数据多要素模型框架；最后基于多要素模型实现对时空大数据的有效组织管理。

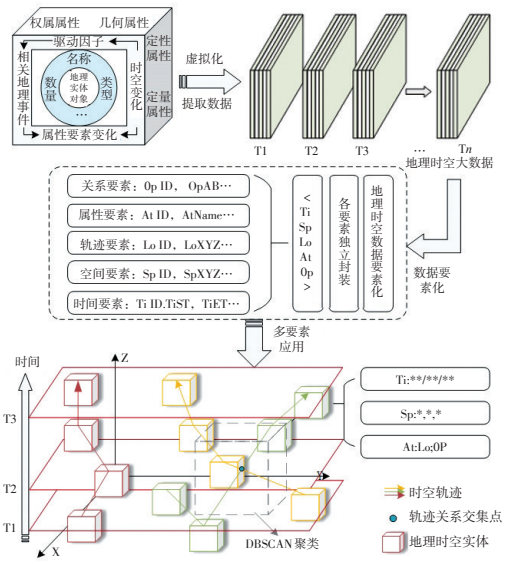


图1　多要素模型框架

3　基于多要素模型的时空大数据组织方法

时空大数据涵盖的信息非常广泛，而人类活动产生的时空大数据80%都与地理信息相关。由于时空大数据具有多源异构和异频采样的特点，其组织方法较

复杂，因此本文基于多要素模型，结合时空大数据的数据结构和存储方式，设计了相应的时空大数据多要素数据库管理方法。

3.1 基于多要素模型的时空大数据数据库设计

时空大数据因其特殊的数据结构和数据关系，需采用空间数据库进行数据存储。地理空间数据库中的地理数据按照一定的数据模型组织、描述和存储，具有较小的冗余度和易扩展性，方便用户共享。由于时空大数据体量庞大，因此基于多要素模型的时空大数据数据库以多要素为基础，将多要素时空大数据统一存储在Oracle Spatial对象—关系数据库中，实现对多要素、空间和时空大数据信息的统一管理，包括现势库、历史库和多要素库（图2）。

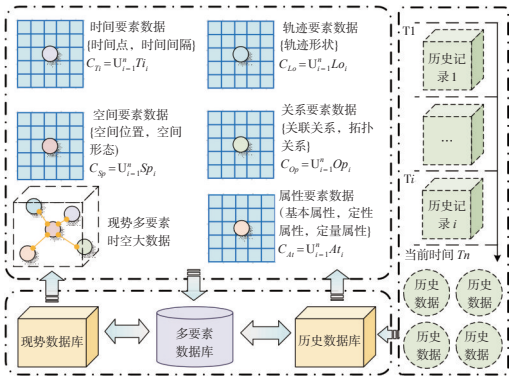


图2 数据库结构设计

1）历史数据库。本文采用一种有利于实现多要素历史数据减量存储的方式，即仅记录同一地理时空对象某一固定回溯时间点的历史数据，其他时间点数据只记录增量或减量数据。设某一固定回溯时间点为 T_0，该时间点下地理时空对象的数据为 HD_0，则当前时间点 T_n 存储的历史数据为：

$$HD = HD_0 + \mathrm{U}_{i=1}^{n}\Delta D_i \tag{5}$$

2）现势数据库存储和管理当前状态下空间对象的最新时空数据，是对当前空间对象的现势性描述[16]。同一地理时空对象在当前时间点仅包含一组时空数据，该组时空数据应只包含一个时间数据。

3）多要素数据库是基于多要素模型的数据库，将现势数据库和历史数据库中存储的时空大数据要素化并进行分要素封装存储，是增、删、改、查时空大数据的关键。针对同一地理时空对象，按照时间、空间、属性、轨迹、关系进行数据要素化，并进行封装标识。单一地理时空对象数据和总数据的存储为：

$$\begin{cases} MD_i = Ti_i + Sp_i + At_i + Lo_i + Op_i \\ MD = \mathrm{U}_{i=1}^{n} Ti_i + \mathrm{U}_{i=1}^{n} Sp_i + \mathrm{U}_{i=1}^{n} At_i + \\ \mathrm{U}_{i=1}^{n} Lo_i + \mathrm{U}_{i=1}^{n} Op_i \end{cases} \tag{6}$$

3.2 基于多要素模型的时空大数据组织方法应用

本文采用微软GeoLift项目的GPS轨迹数据集[17-18]进行应用分析。该数据集收集了2007年4月—2012年8月182个用户的轨迹数据，包括17 621条轨迹，总距离为120多万km，总时间为48 000多h。这些数据不仅记录了用户在家和工作地点的位置轨迹，还记录了大范围的户外活动轨迹，如购物、旅游、远足、骑自行车等。

由于源数据集格式为.plt，因此需要利用Python对数据格式进行预处理。根据要素的表现形式，可将要素分为基础要素和可视化要素两类。预处理后对数据进行基础要素化，即时间、空间与属性要素。其中两组数据的基础要素化结果见表1。

表1 数据基础要素化

要素	标识码	数据1	数据2
时间	TiID	39747.34,8:12:29	39746.21,4:57:55
空间	SpID	(116.3273,39.97746,154)	（116.3208,40.00614,492）
属性	AtID	001,20081023055305	004,20081025045755

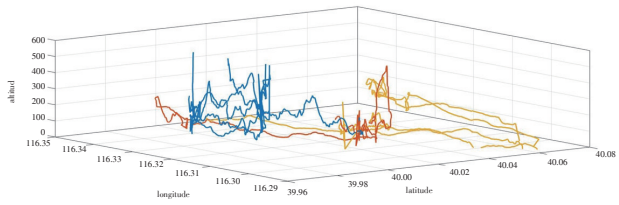
轨迹要素与关系要素则需通过可视化处理才能获取，将预处理后的数据通过Matlab进行三维可视化，一次只使用3组数据进行可视化对照。其中4组数据的轨迹三维图见图3。

根据空间要素相关信息匹配地图，找到时空对象轨迹发生地进行地图可视化和后续分析。利用Map Lab对地图与数据进行融合提炼，获得各组数据在地图上的轨迹和轨迹关系重叠区；再对轨迹关系重叠区的轨迹点进行DBSCAN聚类，结果见图4。根据计算获得的各轨迹点和关联点提取轨迹要素和关系要素，实现更深的数据挖掘，进而挖掘基于多要素模型的时空大数据有效信息。

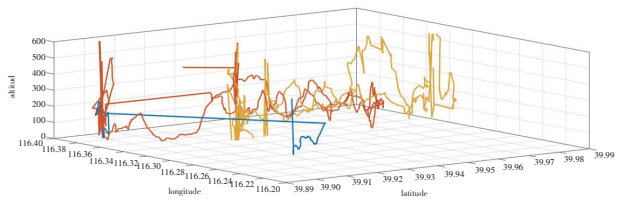
由此可见，在现实生活中利用多要素数据模型可实现对时空大数据的有效组织与管理，并能更深地挖掘时空大数据信息。因此，基于多要素模型的时空大数据组织方法具有进一步研究的意义，也为具有多源异构、动态变化等特征的时空大数据组织与应用提供了新的思路。

4 结 语

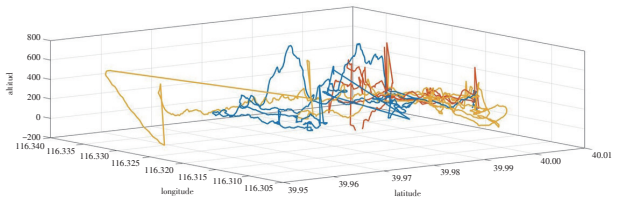
时空大数据研究蕴含着巨大的社会、经济、科研价值，已得到学界的高度重视。在科学技术发展的驱动下，时空大数据的数据量飞速增长，针对时空大数据的大规模数据挖掘、机器学习等新方法也在蓬勃发



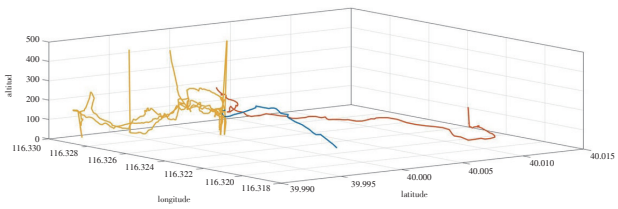
a 第1组三维可视化



b 第2组三维可视化



c 第3组三维可视化



d 第4组三维可视化

图3　基于多要素模型的时空大数据轨迹三维图

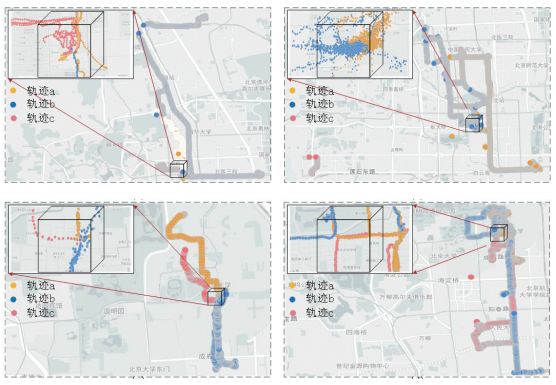


图4　时空大数据轨迹关系可视化

展[19]。智能测绘的兴起为进一步对时空大数据进行组织管理提供了思路。本文通过分析目前时空大数据组织方法的特点，基于时空大数据多要素化的概念，融合时空大数据动态变化的轨迹特征，构建了时空大数据多要素模型；并结合GPS轨迹实例数据进行多要素模型应用分析，验证了时空大数据多要素模型的实用性和有效性，实现了对时空大数据更有效和深入的信息挖掘，为进一步揭示自然物理空间和人类社会活动发展规律提供了参考。

参考文献

[1] 余扬，赵凯飞，沈嘉. 室内定位技术应用、研究现状及展望[J]. 电信网技术，2014，40(5)：46-49

[2] 邓淑丹. 时空大数据全链条管理分析研究与应用[J]. 地理空间信息，2024，22(12)：128-132

[3] Mohamed E，Marco A R. Dynamic Multiscale Spatiotemporal Models for Multivariate Gaussian Data[J]. JASA Journal of the American Statistical Association，2021，41(2)：100475

[4] Yang C，Zhang T Y，Chen Y X，et al. Research and Application Exploration for Spatiotemporal Data Integrated Theory[J]. Journal of Physics：Conference Series，2021(1)：12026

[5] Federal Geographic Data Committee. FGDC Launches New Version of Web-based Geospacial Data[Z]. Sea Technology：Worldwide Information Leader for Marine Business，Science and Engineering，2013，10：50

[6] 刘先林. 为社会进步服务的测绘高新技术[J]. 测绘科学，2019，44(6)：1-15

[7] 龚健雅，张翔，向隆刚，等. 智慧城市综合感知与智能决策的进展及应用[J]. 测绘学报，2019，48(12)：1482-1497

[8] 王家耀. 人工智能赋能时空大数据平台[J]. 无线电工程，2022，52(1)：1-8

[9] 张梦寒，杜德慧，张铭茁，等. 时空轨迹数据驱动的自动驾驶场景元建模方法[J]. 软件学报，2021，32(4)：973-987

[10] 李德仁，马军，邵振峰. 论时空大数据及其应用[J]. 卫星应用，2015(9)：7-11

[11] 王慧，蔡志刚，唐伟，等. 面向时空地理数据库的在线异构整合研究[J]. 地理空间信息，2023，21(1)：69-73

[12] 王育红，刘康晨. 面向新型测绘的时空数据属性连接及扩展方法研究[J]. 测绘通报，2021(5)：5-9

[13] 孟祥武，李瑞昌，张玉洁，等. 基于用户轨迹数据的移动推荐系统研究[J]. 软件学报，2018，29(10)：3111-3133

[14] 李雯. 动态关系下的移动对象位置预测方法研究[D]. 徐州：中国矿业大学，2017

[15] 李景文，邹文娟，田丽亚，等. 基于过程的面向对象时空数据模型数据组织方法[J]. 测绘科学，2013，38(5)：100-102

[16] 黄科佳，李少杰，左尧，等. 基于四叉树索引和双缓存机制的GIS动态可视化模型设计与实现[J]. 测绘与空间地理信息，2018，41(4)：102-105

[17] Yuan H，Li G. A Survey of Traffic Prediction：From Spatio-temporal Data to Intelligent Transportation[J]. Data Science and Engineering，2021，6：63-85

[18] Zheng Y，Xie X，Ma W Y. GeoLife：A Collaborative Social Networking Service Among User[J]. IEEE Data Engineering Bulletin，2010，33：32-39

[19] 朱建军，宋迎春，胡俊，等. 测绘大数据时代数据处理理论面临的挑战与发展[J]. 武汉大学学报(信息科学版)，2021，46(7)：1025-1031