



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860,CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目：因果视角下的地理空间网络表征方法综述：进展、挑战与展望
作者：李海峰，贺丝露，陈海鹏，刘瑜，顾鑫
DOI：10.13203/j.whugis20240366
收稿日期：2024-12-22
网络首发日期：2025-01-14
引用格式：李海峰，贺丝露，陈海鹏，刘瑜，顾鑫. 因果视角下的地理空间网络表征方法综述：进展、挑战与展望[J/OL]. 武汉大学学报(信息科学版).
<https://doi.org/10.13203/j.whugis20240366>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20240366

引用格式：

李海峰, 贺丝露, 陈海鹏, 等. 因果视角下的地理空间网络表征方法综述：进展、挑战与展望[J]. 武汉大学学报（信息科学版）, 2025, DOI:10.13203/J.whugis20240366 (LI Haifeng, HE Silu, CHEN Haipeng, et al. Survey on Geospatial Network Representation Learning from a Causal Perspective: Advances, Challenges, and Prospects[J]. Geomatics and Information Science of Wuhan University, 2025, DOI:10.13203/J.whugis20240366)

因果视角下的地理空间网络表征方法综述：进展、挑战与展望

李海峰¹, 贺丝露¹, 陈海鹏^{*2}, 刘瑜³, 顾鑫²

1 中南大学地球科学与信息物理学院, 湖南 长沙, 410083

2 中国运载火箭技术研究院, 北京, 100076

3 北京大学地球与空间科学学院, 北京, 100871

摘要：地理空间网络（geospatial network）是指网络中的节点或边的属性与地理位置相关的网络，能够同时表达地理研究中的地理实体及实体间的关系，是GIS研究中重要的一种数据类型。随着深度表征学习的发展，利用深度模型自动提取地理空间网络表征逐渐成为当前主流的模式。现有表征学习范式通过拟合观测数据中的相关性作为可学习的信号，但无法处理地理空间网络的某些特性带来的挑战。因果表征学习旨在为现有表征学习范式进一步引入因果知识作为学习约束，具有解决这类挑战的潜力。论述了为地理空间网络表征引入因果性视角的必要性，并对地理空间网络因果表征学习现有进展进行了总结。为了更好地提炼地理空间网络因果表征学习的重要概念及基本问题，以因果学习中的常用框架——结构因果模型（Structural Causal Model, SCM）为主线，将地理空间网络因果学习形式化为三个子任务。同时，构建了地理空间网络表征学习的统一框架，基于该框架总结了引入因果学习策略的意义。将地理空间网络中的因果关系分为空间因果和时间因果两类，并对现有研究分别进行了总结和梳理。在此基础上，对现有研究存在的挑战进行了总结，同时探讨了未来研究的方向。除此之外，还对通用的地理空间网络表征学习框架进行了抽象，为地理空间网络研究引入其他假设提供了可参考的骨架。

关键词：地理空间网络表征学习；因果学习；地理空间因果；综述

Survey on Geospatial Network Representation Learning from a Causal Perspective: Advances, Challenges, and Prospects

LI Haifeng¹, HE Silu¹, CHEN Haipeng^{*2}, LIU Yu³, GU Xin²

1 School of Geosciences and Info-physics, Central South University, Changsha 410083, China

2 China Academy of Launch Vehicle Technology, Beijing 100076, China

3 School of Earth and Space Sciences, Peking University, Beijing 100871, China

Abstract: A geospatial network refers to a graph whose nodes or edges are associated with geographic locations. As an important data type in GIS, it simultaneously models geographical entities and their relationships. With the development of deep representation learning, leveraging deep models to automatically extract representations of geospatial networks has become the mainstream paradigm. The

收稿日期：2024-12-22

基金项目：国家自然科学基金 42271481；湖南省自然科学基金 2022JJ30698

第一作者：李海峰，博士，教授，博导，现从事多模态时空大模型，时空因果，时空智能体方向研究。

lihaifeng@csu.edu.cn

通讯作者：陈海鹏，硕士，研究员。chenhpcalt@163.com

existing representation learning paradigm adopts the correlations observed in data as learnable signals but is unable to address the challenges posed by certain characteristics of geospatial networks. Causal representation learning aims to introduce causal knowledge as a learning constraint within the existing paradigm, offering potential solutions to such challenges. This article discusses the necessity of introducing causal assumptions for geospatial network representation and summarizes the current progress in causal representation learning for geospatial networks. To better distill the key concepts and fundamental issues of causal representation learning for geospatial networks, this article uses one of the most frequently used frameworks, the structural causal model, as the common thread, formalizing causal learning in geospatial networks into three sub-tasks. A unified framework for geospatial network representation learning is constructed, based on which the significance of introducing causal learning strategies is summarized. Additionally, this article categorizes causal relationships in geospatial networks into spatial causality and temporal causality, providing a comprehensive review and analysis of existing studies in each area. The challenges in current research are summarized, and future research directions are explored. Furthermore, this article abstracts a general framework for geospatial network representation learning, offering a reference skeleton for introducing other assumptions into geospatial network research.

Key words: Geospatial Network Representation Learning; Causal Learning; Geospatial Causality; Survey

如“任何事物都与其他事物相关”所述，地理学第一定律^[1]揭示了地理单元之间普遍存在着联系，这种区域联系是探索地理时空过程的重要视角之一^[2]。作为 GIS 最早出现也最重要的一种数据表示，网络（network）能够同时表达地理空间研究中的一阶量（即地理单元及其对应的属性）和高阶量（即地理单元间的关系）^[3]，灵活建模不规则的地理单元分布，引导了 GIS 数据结构的革命性进步^[4]。在该视角下，地理空间数据被组织为网络结构，使用节点表示特定的地理空间对象，边则用于表达对象间的关系，这类数据被称为地理空间网络（geospatial network）。地理空间网络可以视为空间网络的一个特殊类型，包括所有节点或边的属性与地理位置相关的网络^[5]，如路网、基于位置的社交网络等，特性在于能够刻画地理对象间的静态和动态关系，表达地理研究中关注的空间分布和空间交互^[6,7]等问题，因而成为地理学及相关领域中的重要研究对象。随着地理大数据时代的来临，地理空间网络数据及其蕴含的知识日益丰富，以地理空间网络为研究对象，从具有全面、真实反映地理各要素的状态以及演化过程潜力的大数据^[8]中进行智能理解和知识发现，对于理解地理空间格局^[9]和人类活动行为^[10]至关重要，而地理空间人工智能（Geospatial Artificial Intelligence, GeoAI）的出现使其成为可能^[11,12]。

GeoAI 领域算法的成功高度依赖于数据特征，而表征学习避免了有限的、高成本的手工特征工程^[13]，利用神经网络模型为数据样本自动生成实值向量进行表示^[14]，是当前地理空间网络智能学习的主流范式。其中，图卷积网络^[15]（Graph Convolution Network, GCN）等图表征模型表现出了出色的地理空间网络表征能力，如 GCN 被用于学习推特用户地理定位网络^[16]和城市场所网络^[17]的表征。这类模型的表征能力不仅仅得益于模型架构，更需要从观测数据中捕获某种可学习的模式作为信号，现有表征模型主要是将从观测数据中拟合的相关性作为待学习的模式。然而，从地理空间网络观测数据中学习这类正确的信号面临诸多挑战，原因在于地理空间网络数据作为地理空间数据的一种，从生成到获取过程中受到地理过程的复杂性、地理环境的差异性和观测手段的有偏性的影响，使得数据中原本存在的模式难以直接被发现，需要对数据生成过程中的影响因子和环境因子进行刻画，并去除数据获取过程中干扰因子的影响。

因果性作为数据中严格的一种假设约束，具有解决上述问题的潜力。而基于因果性假设衍生的因果学习作为一类有效的数据分析手段，则提供了一系列可执行的学习方法，能够在实践中解决上述问题。本文沿用数据因果研究中的一种观点：因果可以视为塑造数据生成的一种力量^[18]，即因果是指数据生成过程中影响因子之间的相互作用，对地理空间数据表征存在的问题进行描述。如图 1，因果发现方法能够构建数据生成过程中影响因子的因果关系，并将环境因子作为外生变量建模在因果机制中，同时，因果推断方法提供了系统的去混杂策略去除干扰因子的影响。因此，为地理空间网络数据表征学习引入因果性假设和因果学习策略，能够更好地揭示地理空间网络数据背后的生成机制，实现地理空间网络数据理解和知识发现。本综述主要探讨这类地理空间网络因果表征学习的相关工作。

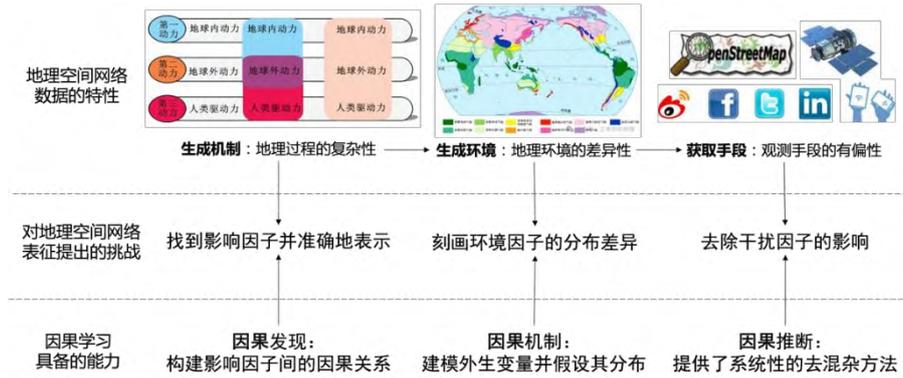


图 1 地理空间网络表征学习面临的挑战及因果学习具备的能力

Fig.1 Challenges in Geospatial Network Representation Learning and the Capabilities of Causal Learning

为了更清晰地组织地理空间网络因果表征的基本问题，本文先分别对地理空间网络表征（第 1 节）和因果学习（第 2 节）进行了关键问题及统一框架的抽象，并基于此对现有地理空间网络因果表征学习方法（第 3 节）的进展进行了梳理和介绍。具体来说，本文基于地理空间网络表征过程中的两个不同层级的概念，对地理空间网络因果进行了分类；基于因果学习中的结构因果模型，将地理空间网络因果学习形式化为 3 个子任务。最后，因果表征学习在地理空间网络上的应用仍处于初步探索的阶段，面临许多本领域特有的挑战，本文在第 4 节探讨了现有的挑战，并展望了潜在有效的解决方式。本文综述的整体组织逻辑与结构如图 2：

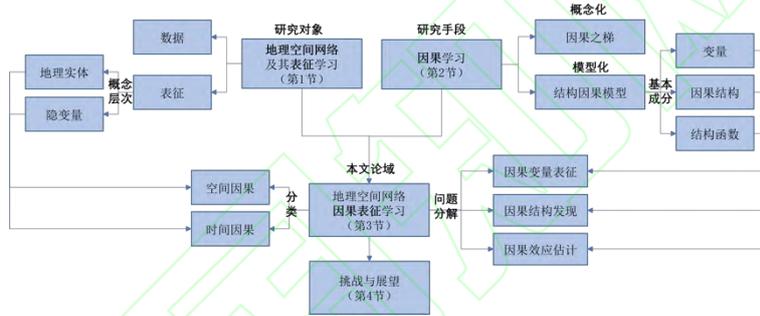


图 2 综述整体组织逻辑与结构

Fig.2 Overall Organizational Logic and Structure of the Review

1 地理空间网络及其表征

地理空间网络是本综述面向的研究对象。目前对于地理空间网络的定义和形式化表达，以及各领域地理空间网络数据及其表征尚未有统一的框架进行总结和标准化。因此，本节首先对地理空间网络进行了定义及介绍，并抽象出了地理空间网络表征的统一框架，该框架适用于所有的地理空间网络表征方法及应用。

1.1 地理空间网络

地理空间网络是指网络中的节点或边能够与地理位置相关联的网络^[5]，包含不同领域的多种数据集。常见的地理空间网络有交通运输网络、社交网络、人类移动网络、疾病传播网络、气候网络、贸易网络和城市网络。交通运输网络包括高速公路数据集 METR_LA、Los-loop 和 PEMS_BAY，城市路网数据集 SZ-taxi、TAXIBJ 和 NYCTAXI，目前也有公开的基准库^[19]；社交网络特指基于位置的社交网络^[20]（Location-based Social Network, LBSN），如用户注册数据（包含注册地）、微博打卡数据和美团打卡数据等；人类移动网络建模地区间的人口移动，小至场所，大至国家，包括航空网络、出发地-到达地（Origin-Destination, OD）网络等，航空网络^[21,22]从机场和航线中构建而来，可以刻画航空服务的客流量、直航距离和单程票价等特征，OD 网络则可以通过地铁刷卡数据、出租车行程数据构建；疾病传播网络^[23]背后也隐含着人类的移动，但是更注重人类移动对疾病传播的影响；气候网络^[24,25]刻画研究区域内局部的气候分布，如海平面、气象观测站等；贸易网络^[26]刻画地区间以贸易为基础的交互；城市网络则刻画城市尺度的局域连续，如城市功能网络^[17]通常以社区或场所等城市功能的基本单元为节点，刻画节点之间的交互，居民地网络^[27]则通过居民地之间的交通通达关系构建。除此之外，还存在一种特殊的地理空间网络类型，这类地理空间网络包含了节点间不同类型的关系，即地理多元流网络^[28]。在全球尺度上，有融合贸易、投资、人类移动、知识流和外交等多种关系的国际关系多元网络^[29]和基于媒体、立法执法和叛乱事件构建的国际关系多元网络^[30]。在城市及城市群尺度上，有基于人流、物流、经济流和技术流构建的城市经济圈空间网络^[31]，有基于不同交通方式构建的城市多元交通网络^[32]，有融合了城市间金融、信息和交通联系的城市群多元网络^[33]

等。综上所述，地理空间网络覆盖了多种不同的领域和应用，无论作为空间网络数据还是地理数据的子类来说都具有重要的研究与应用意义。

1.2 地理空间网络表征框架

1.2.1 地理空间网络数据

广义的地理空间网络包括所有节点或边具有地理位置属性的网络数据^[5]，主要建模地理实体的分布及其之间的关联关系。其中地理实体指通过观测特征描述的地理空间中存在的对象，如事件、地理区域、空间分布的传感器和行人等，往往包含空间、时间和属性等特征；当观测到一个以上的地理实体时，不同地理实体间的关联关系可以通过不同的视角构建，如时间视角下的接续关系、空间视角下的邻近关系和属性视角下的相似关系等。地理空间网络可以形式化表示为一张图 $G = (V, E, F)$ ，其中节点为地理实体 GE ，边则代表地理实体间的关联关系 R 。节点集 $V = (GE_1, GE_2, \dots, GE_n)$ ；边集 E 中的元素 $e_{ij} = R(GE_i | GE_j)$ 代表节点 i 和 j 间的特定关系，通过预定义规则得到， E 可以进一步构建为邻接矩阵 $A_{n \times n}$ ；特征集 $F_{n \times m}$ 代表节点的特征值， n 代表节点数量， m 代表节点的特征维度，往往从观测特征中构建得到。

1.2.2 表征模型

鉴于地理空间网络数据的非欧结构，地理空间网络通常被表示为图。因此，现有的图表征方法潜在地可以作为地理空间网络表征的方法。地理空间网络表征学习与地理空间数据表征学习的最大区别在于顾及地理实体之间的关联关系，因此需要考虑一个关键问题：如何学习保留图结构的图表征？本文从图结构建模的角度对现有图表征模型进行考察，将现有骨干模型分为图嵌入、图神经网络、图 Transformer 和大语言模型四种类别，分类见表 1。

表 1 地理空间网络表征模型分类

Tab.1 Taxonomy of Geospatial Network Representation Models

类别	基本策略	子类别	现有方法
图嵌入	显式地保留原始图拓扑结构，得到低维嵌入	基于微观拓扑结构	DeepWalk ^[34] , Node2Vec ^[35] , High-order Proximity Preserved Embedding (HOPE) ^[36]
		基于中观拓扑结构	Modularized Nonnegative Matrix Factorization (MNMF) ^[37] , 基于模块最大化 ^[38] 、谱聚类 ^[39] 、边聚类 ^[40] 的社区检测算法，考虑结构相似性的 Struct2Vec ^[41] 和 GraphWave ^[42]
		基于宏观拓扑结构	考虑惩罚节点度最大的节点间的邻近性 ^[43] , Hierarchical Representation Learning for Networks (HARP) ^[44]
图神经网络	隐式地保证特征变换过程中图结构的不变性和等变性，利用卷积操作实现局部特征的聚合	基于谱的卷积	ChebNet ^[45] , GCN ^[15] , AGCN ^[46] , DGCN ^[47] , SCNN ^[48]
		基于空间的卷积	DCNN ^[49] , GraphSAGE ^[50] , GAT ^[51]
图 Transformer	利用自注意力捕捉潜在图结构的同时，通过图结构和位置来保留原始图结构	位置编码或（和）结构编码	GT ^[52] , SAN ^[53] , DGT ^[54] , Graphormer ^[55] , GraphiT ^[56] , GOAT ^[57] , GRIT ^[58]
大语言模型	将图结构转换为大语言模型可以理解的序列，利用大语言模型的综合推理能力得到图的表征或者直接生成下游任务的结果	基于图描述语言	直接生成问题答案 ^[59-63]
		基于图结构编码	GraphLLM ^[59]

1.2.3 下游任务

地理空间网络表征能够应用于多种下游任务，表征在下游任务中的表现可以用于检验学习到的表征，理想的地理空间网络表征能够在下游任务中具有良好的表现。根据任务对象在图结构中的层次可以将下游任务分为节点级别，边级别和图级别。

(1) 节点级别。节点级别的地理空间网络表征下游任务旨在为网络中的每个节点得到相应的任务输出，比较典型的包括预测任务和分类任务。其中，预测任务利用每个节点的历史时序观测数据得到未来时间步的观测数据，如交通流预测^[64]，犯罪预测^[65]，气候预测^[66]，空气质量预测^[24]和海平面预测^[25]等；分类

任务旨在为每个节点得到感兴趣的类别标签，如车辆行为分类^[67]和场所功能分类^[17]。

(2) **边级别**。边级别的地理空间网络表征下游任务旨在预测网络中两个节点之间是否存在连边或边的属性值，较典型的应用是 OD 流预测^[68, 69]。

(3) **图级别**。图级别的地理空间网络表征下游任务旨在为整个或局部的地理空间网络得到相应的任务输出，包括子图级别和整图级别的任务。子图级别如社区检测^[70]，为不同子图学习到标签；整图级别如气候变化模式探测^[71]，将城市多年来的气候特征组织成图，将城市气候模式分类为稳定/不稳定。

1.2.4 表征框架

基于上述定义，地理空间网络表征的过程可以描述为：从时空观测中构建得到地理网络数据 $G = (V, E, F)$ ，地理空间网络表征模型在假设约束下通过学习得到 G 在表征空间中的嵌入（见式 (1)）。不同于原始输入中 F 表示的观测变量，表征学习能够学习地理实体在不可观测的隐变量空间中的表示。理想的隐变量空间能够描述数据背后高层的、抽象的解释因子，假设最终得到的表征空间共 k 维，每个隐变量可以用不同的几个维度的特征进行表示。

$$Embedding(G) = F_{ebd}(G, hypothesis) = \{H_{i1}, H_{i2}, \dots, H_{ik}\}_{i=1, \dots, n} \quad (1)$$

其中 $F_{ebd}(\cdot)$ 为地理空间网络表征模型。将 $Embedding(G)$ 输入到下游任务执行模型中，得到下游任务的结果（见式 (2)）。其中 $F_{downstream}(\cdot)$ 为下游任务执行器，可以为分类，回归，聚类等模型， \hat{Y} 为下游任务标签。

$$\hat{Y} = F_{downstream}(Embedding(G)) \quad (2)$$

综上所述，地理空间网络表征框架如图 3。需要强调的是，由于表征学习的一个重要假设是认为学习到的特征表示能够揭示数据背后隐含的某种知识或模式，因此，在框架图中可以添加由表征指向数据的箭头，从而形成“数据→表征→数据”的信息流动闭环。

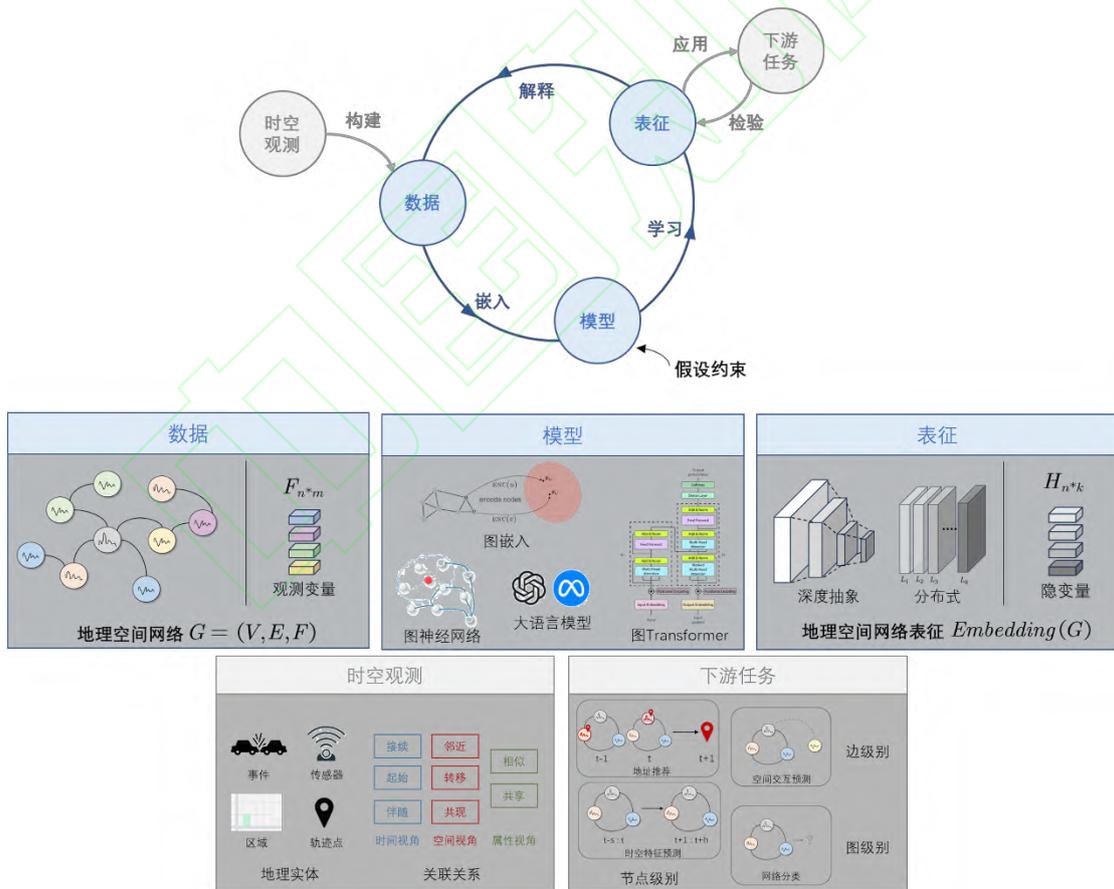


图 3 地理空间网络表征框架图

Fig.3 Framework of Geospatial Network Representation

2 因果学习

因果学习为地理空间网络表征过程中引入因果知识提供了系统性的理论和方法作为引导。本节首先从概念上介绍了因果性与相关性的差异，然后以因果学习中的一种统一语言框架——结构因果模型 (Structural

Causal Model, SCM) 作为统一的主线, 对因果学习中的关键问题进行了抽象, 并导出对地理空间网络因果学习的形式化。

2.1 攀登因果之梯：从相关到因果

因果之梯是朱迪亚·珀尔在对于理解因果关系的三个层级的抽象, 并分别对应着学习者的三种不同层次的认知能力的划分^[18]。寻找和理解真正的因果关系, 正是逐层攀登因果之梯的过程, 如果你从未到达低的一层, 那就无法抵达更高的一层。

第一层是关联。在这一层级的主要行为是通过观察寻找规律, 这种规律能够指示某种关联性。在统计学中, 这种关联性可以通过条件概率来表示。

第二层是干预。干预不再受困于被动收集的数据, 而是通过创造某种控制条件的场景并从中获得新的数据。因此, 干预层能够回答的问题是: 如果改变某一个对象的状态, 另一个对象会如何变化, 这种关系无法直接通过条件概率来表示, 需要使用因果模型中的干预算子才能表示。

第三层是反事实。反事实不仅仅停留在对现有世界的干预上, 而是需要跳出事实, 去想象回到过去的某个节点并改变历史, 在虚拟世界中会对应着什么样的结果。在这种想象思维中, 需要去寻求背后解释的推断, 这就需要人类最高级的认知能力, 对应着因果之梯的最高层。

现有的大部分深度表征模型由于依赖于学习训练数据中的模式而没有考虑因果知识的建模, 停留在因果之梯的底层, 抵达干预层意味着需要在模型训练过程中改变其训练的环境规则。本文所探讨的因果视角下的地理空间网络表征学习方法, 则是利用了某种关于因果的假设, 并设计了干预性或观察性的方式模拟干预层级甚至反事实层级的情景, 从中学习并建模因果知识, 从而使得表征能力能够从第一层攀登到第二层及更高层。

2.2 结构因果模型

SCM 是因果相关研究中表示因果机制的一种常见框架, 换言之, 可以理解为描述因果的一种语言。在所研究的系统内部, 变量之间相互影响, 在自然赋值和物理机制的作用下呈现出其本身的状态, SCM 能够描述这类变量通过因果机制得到系统观测值的过程。为了表示系统中感兴趣的变量及其之间的关系, SCM 使用了一系列的结构等式 $X_i = f_i(PA_i, U_i)$, $i = 1, 2, 3, \dots, n$ 。其中, X_i 为内生变量, PA_i 为父母变量, 代表直接影响变量 X 取值的变量集; 每个变量 X_i 都对应一个外生变量 U_i , 代表由系统外部的因子造成的扰动, 所有外生变量之间彼此独立^[72]。对于每个变量 X_i 都存在一个结构等式 f_i , 用于描述该变量的父母节点和噪声变量如何函数化地影响该变量, 这种函数化的关系可以理解为一种自然的物理法则, 即当为父母变量和噪声变量赋值时, 变量 X 在自然状态下会得到相应的取值^[73]。

每一个 SCM 都可以关联到一个图形化的因果模型作为图形化的表示, 称为因果图。其中, 内生和外生变量用图节点表示, 变量间的因果关系用边表示, 由原因变量指向结果变量, 结构等式 f_i 可以通过指向变量 X_i 的边的集合来表示。如图 4, 左侧的图为右侧的 SCM 对应的因果图, 基于 SCM 的因果学习可以直接通过在因果图的操作来实现, 并且提供了相应的方法实现对系统状态的估计。

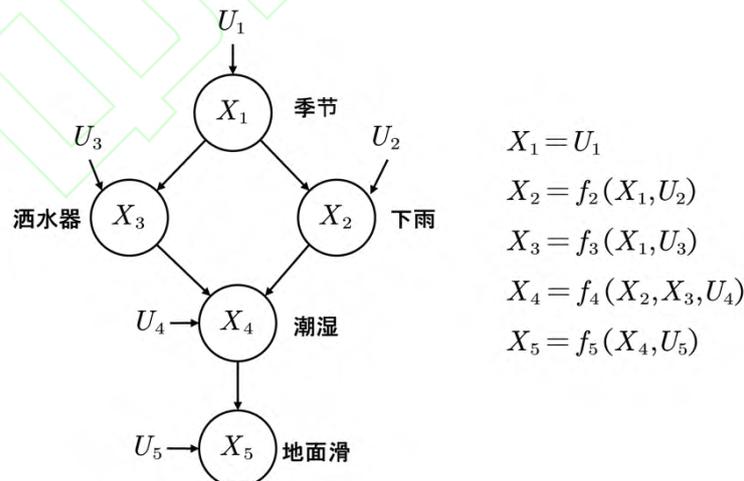


图 4 SCM (右) 及其对应的因果图 (左)

Fig.4 SCM (Right) and The Corresponding Causal Diagram (Left)

从整体上来看, SCM 建模的信息进一步可以被抽象为三类, 变量, 结构和结构函数。相应地, 因果图通过三个成分来表示这三类信息。为了建模地理空间网络背后的 SCM, 地理空间网络因果学习需要设计相应的子任务对三个成分进行表达, 因而可以导出地理空间网络因果学习的三个子任务。基于 SCM 框架对地理空间网络因果学习流程进行形式化的逻辑关系如表 2。

表 2 SCM、因果图和地理空间网络因果表征子任务的对应关系

Tab.2 Correspondence between SCM, Causal Diagram and Subtasks of Geospatial Network Causal Representation

SCM	因果图	地理空间网络因果表征子任务
变量	节点	因果变量表征
因果结构	边结构	因果结构发现
结构函数	边系数	因果效应估计

作为一种系统化的因果建模语言，SCM 能够实现对地理空间网络背后因果机制的统一和完整表达，建模各因果变量及其之间的因果关系。以水文网络为例，上游河流的水流量作为原因变量影响着下游河流的水流量，SCM 能够通过因果图建模整个河网系统的底层因果机制，将系统中的所有河段表示为节点，河段间的上下游关系表示为边结构，并根据原因节点引起的结果节点水流量变化的大小得到边系数。同时，现有因果学习的研究者基于 SCM 定义了一系列因果发现 [74] 算法和因果推断 [75] 算法，能够分别完成地理空间网络上的因果结构发现和因果效应估计任务。

2.3 地理空间网络因果学习

为了保证整个因果学习框架的系统性与合理性，本文选择了因果学习研究领域公认的一种标准语言——SCM，作为统一因果学习框架的骨架，并相应地将地理空间网络因果学习的过程分解为三个子流程。由这三个子任务组成的地理空间网络因果学习流程如图 5。

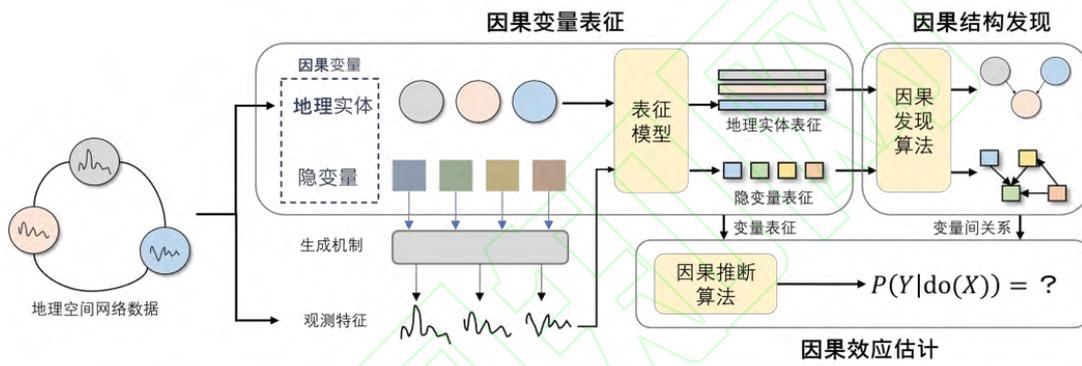


图 5 地理空间网络因果学习流程

Fig.5 Pipeline of Geospatial Network Causality Learning

2.3.1 因果变量表征

因果变量表征是指从数据中为感兴趣的因果变量学习到抽象的、表达性的表示的过程。根据不同类型的观测数据，可以引入不同结构的模型获得表征。例如，当观测数据为图像时，可以利用变分自编码器作为因果变量的表征模型 [76]；当观测数据为文本时，可以利用语言模型（如 BERT）作为因果变量的表征模型 [77]；当观测数据为时空图时，则可以利用时空图神经网络作为因果变量的表征模型 [78]。由于地理空间网络数据主要为时空数据，具有丰富的时态特征，因此循环神经网络（Recurrent Neural Network, RNN）和卷积神经网络（Convolutional Neural Networks, CNN）常被用于从时序数据中提取因果变量的表征，例如长短时记忆网络（Long Short Term Memory Network, LSTM），门控循环单元（Gated Recurrent Unit, GRU）和时域卷积网络（Temporal Convolutional Network, TCN）。只有当感兴趣的因果变量被表示为表征向量后，才能进行后续的因果结构发现和因果效应估计任务，部分因果学习算法支持因果变量的原始观测特征作为输入，则可以将结构化后的观测特征视为因果变量的初始表征。

2.3.2 因果结构发现

因果结构发现是指利用因果发现算法学习出因果变量之间的因果关系，旨在发现变量间更本质的依赖关系，进而能够利用已知变量建立对未知变量的相关信息。根据对于因果机制的假设和数据类型的不同，可以选择不同的因果发现方法进行因果结构的发现。如格兰杰因果检验算法等算法可用于时序因果发现 [79]，还有一系列方法可用于非时序数据的因果发现 [80]。

2.3.3 因果效应估计

因果效应估计是指利用已知的因果关系，利用因果推断方法计算变量之间的因果效应，进而能够得知原因变量发生变化时，结果变量会如何产生相应的变化。根据因果结构和观测数据的不同，可以使用不同的因果推断方法 [81] 进行因果效应的估计，如前门调整和后门调整等。

3 地理空间网络因果表征学习进展

地理空间网络因果表征学习涵盖了所有在地理空间网络表征的过程中引入了因果知识作为学习约束的工作。为了更好地提炼基本问题和梳理已有进展，需要回答两个重要的问题：什么是地理空间网络因果？

以及如何学习地理空间网络因果？为了回答这两个问题，本节对地理空间网络因果的两种类别进行了划分与定义（3.1.1 节），并提供了将地理空间网络因果学习融入通用表征框架的基本范式（3.1.2 节），在此框架下总结了引入因果学习对于地理空间网络表征的意义（3.1.3 节）。最后，本节对现有相关工作进行了介绍，为了简洁直观地对现有地理空间网络因果表征相关工作中的表征模块（1.3 节）和因果学习模块（2.3 节）进行介绍，本文分别将空间因果和时间因果相关工作总结在表 3 和表 4 中。

3.1 地理空间网络因果表征学习

3.1.1 地理空间网络因果的两种类型

由图 3 可知，地理空间网络表征具有两个基本对象——地理实体及其之间的关联关系。如图 6，经过表征模型的学习后，地理实体可以通过不同的隐变量进行描述。隐变量是隐藏或潜在的变量，是具有实际意义但无法被直接观测的变量^[82]，在本文中特指用于描述地理实体抽象特征的变量；而关联关系通过不同的地理实体间的关系进行描述。

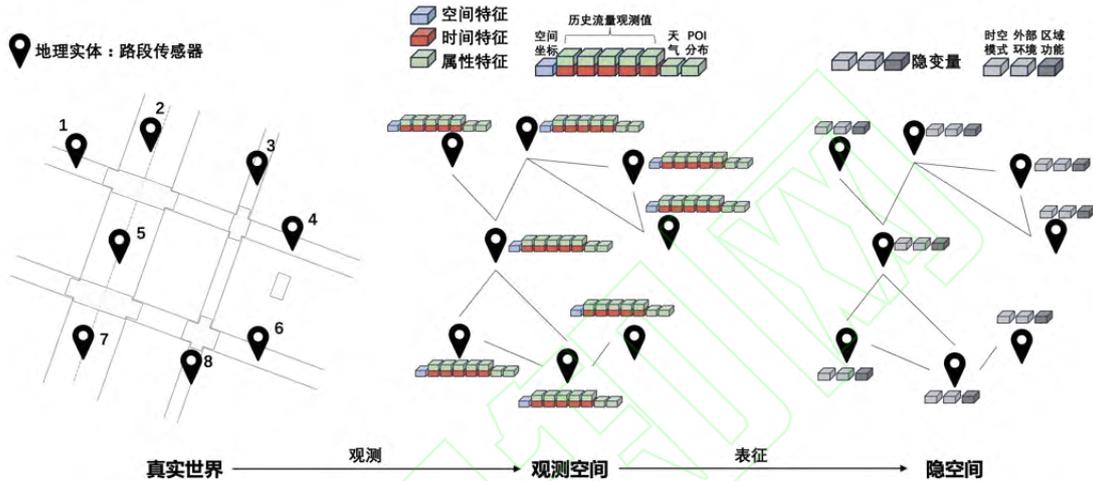


图 6 地理实体在不同空间中的表示

Fig.6 Representations of Geographical Entity in Different Spaces

地理空间网络表征的本质在于通过学习上述两个基本对象的表示，揭示地理空间网络数据背后隐含的机制。如图 7 所示，即使在统一的地理空间网络表征框架（1.2.4 节）下，因果性视角与传统的相关性视角具有本质的不同。由于因果表征学习试图用因果模型逼近数据背后的生成机制，在表征地理实体时，因果表征模型用具有因果关系的隐变量对地理实体进行表示，这种因果关系描述的意义是数据生成过程中，影响因子间如何随时间发生相互作用；在表征地理实体间的关联关系时，因果表征模型将传递了某种影响的地理实体视为具有因果关系，这种因果关系描述的意义是不同的地理实体间如何通过跨空间交互产生影响。

地理空间网络表征模型		相关性视角	因果性视角
目标		揭示地理空间网络背后的隐含机制	
基本假设		数据 $\xrightarrow{\text{拟合}}$ 机制 $\xleftarrow{\text{代理}}$ 模型	模型 $\xrightarrow{\text{推断}}$ 机制 $\xrightarrow{\text{生成}}$ 数据
表征的基本对象	节点 (地理实体) $GE = \{E_1, E_2, \dots, E_k\}$	 隐变量相互独立 $E_1 \perp E_2 \perp E_3 \perp \dots \perp E_k$	 隐变量因果关联 $E_k = f_k(PA_k, U_k), PA_k \subset E$
	边 (关联关系) $R(GE_i GE_j)$	 $IF R(GE_i GE_j) > \epsilon,$ $\epsilon = ST \text{ correlation}$	 $IF R(GE_i GE_j) > \epsilon,$ $\epsilon = \text{Interactive influence}$

图 7 相关性视角和因果性视角下地理空间网络表征的区别

Fig.7 Distinctions Between Geospatial Network Representation from Correlation and Causal Perspective

由于因果表征学习的核心在于发现和解释数据中变量的因果关系，根据这两类因果关系所描述的对象的不同，可以将地理空间网络因果中的因果变量分为地理实体和隐变量两种。进一步地，以这两种变量间的因果机制发生作用时的主要场所作为区分，可以基于此将地理空间网络因果分为以下两类：

(1) **空间因果**。由于受到空间分布的约束，地理实体发生相互作用时需要跨越空间才能完成交互。因

此，描述地理实体间关系的因果机制的发生作用时的主要场所是空间，本文将这类地理空间网络因果称为空间因果。

(2) **时间因果**。隐变量及其之间的关系对于每一个地理实体都适用，即隐变量间的因果机制具有空间不变性，应当在所有地理实体上都满足。由于对于一个空间实体来说，隐变量发生相互作用的过程主要消耗时间，我们将这类地理网络因果称为时间因果。

这两类因果所描述的物理意义也不相同。空间因果关系可以理解为一个实体对另一个实体产生了某种可传递的影响，例如：在交通路网中，由于车流的移动，路段 A 的拥堵会导致路段 B 的拥堵。其中，路段 A 和 B 都可以视为具体存在的地理实体；而时间因果关系则是抽象出高层影响因子之间的作用机制，例如：在出行轨迹预测模型的工作过程中，历史出行模式是外推出未来出行模式的原因，而天气则作为同时影响两者的混杂变量。其中，历史和未来出行模式、天气都可以视为隐变量，是抽象后的概念。本节对两类因果中的现有进展分别进行介绍，若有工作同时考虑两类因果，本文则分别对工作中对空间因果和时间因果的处理进行讨论。

3.1.2 地理空间网络因果表征学习框架

地理空间网络因果学习模块通过因果变量表征、因果结构发现和因果效应估计三个子任务中的至少一个，用于建模从地理空间网络观测数据中学习到的隐含因果知识。这种知识可以表示为因果图、因果效应的强度等不同形式，但最终都需要通过某种方式融入到地理空间网络表征模型中，才能获得因果性的表征。基于 3.1.1 节中的地理空间网络因果的两种类型，与 1.2.4 节中地理空间网络表征的统一框架相结合，可以将地理空间网络因果表征学习框架绘制如图 8。从整体框架来看，在地理空间网络表征学习框架中引入因果学习策略的目的在于捕获地理实体和隐变量层级的因果机制，实现对地理空间网络数据背后生成过程的理解和建模，这种关于因果机制的知识能够融入到地理空间网络表征模型中，作为假设约束引导学习地理空间网络表征。同时，地理空间网络因果学习模块不仅仅能够揭示和解释研究者感兴趣的机理，并且能够在此基础上通过机理的外推得到新的信息。

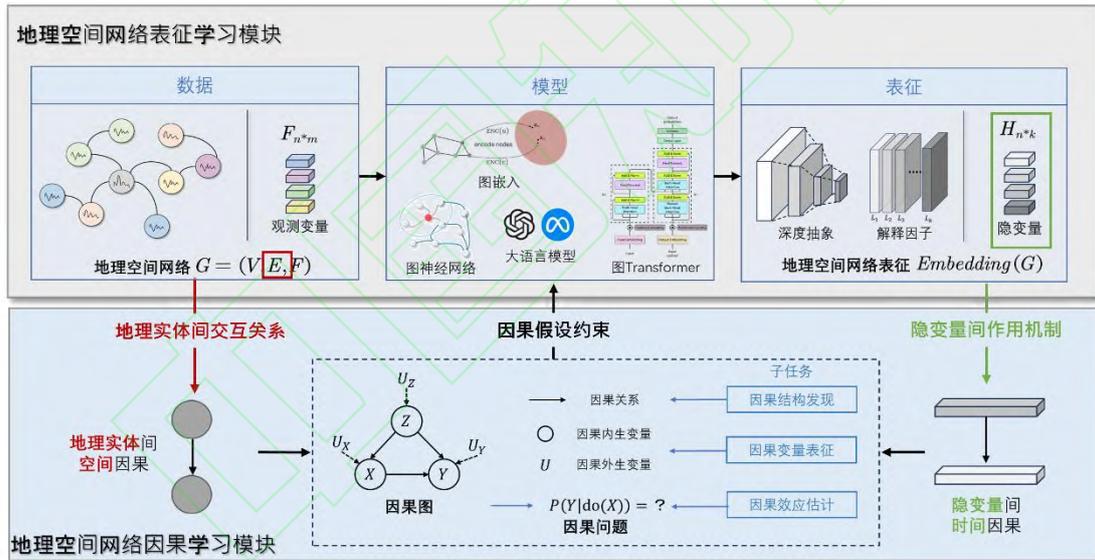


图 8 地理空间网络因果表征学习框架

Fig.8 Overall Framework of Geospatial Network Causal Representation Learning

基于该学习框架，本文分析和总结了引入因果学习对地理空间网络表征的意义，分别如下：

(1) **实现关系挖掘和知识发现**。如图 7 所示，从表征的基本对象的视角来看，空间因果能够捕获不同地理实体之间的因果关系，这种因果关系不同于一般的时空或属性域定义的关联关系，而是能识别发生了跨空间信息流动的实体。这不仅能够挖掘仅从时空或属性值上无法直接推测的地理实体间的关系，还有潜力捕获大尺度下远距离地理实体间的隐含关联。例如 STGC-GNNs^[78]识别到交通路网中具有长程空间依赖关系的路段，路段上分布的测速传感器间的空间距离非常远，但是分别处于同一条高速公路主干道的上下游，因此必然存在着以车辆为载体的信息的流动，这是仅凭“相近的事物更相关”的地理学第一定律无法捕捉到的关系。时间因果则能够揭示高层的隐变量间的相互作用关系，这有利于实现对不同地理概念间关系的知识抽取，并且可以通过构建知识图谱等方式完成地理知识推理。此外，研究者们可以根据不同的地理场景借助专家知识对因果假设进行不同程度的松弛，构建具有不同可靠性的因果关系图。借助因果学习挖掘到的关系和知识能够作为约束项引导地理空间网络表征模型学习到符合预设前提的表征。

(2) **提升地理空间网络表征的可解释性**。现有主流的地理空间网络表征模型大部分都是深度神经网络

络架构, 通过大量参数拟合输入数据和输出结果之间的复杂映射关系获得良好表现。然而, 这种“黑箱”模型中的特征映射过程无法解释为人类可以理解的作用规律, 中间过程中获得的高维表征也难以解译为自然语言可以描述的具体对象。因果学习为提升表征的可解释性提供了有效手段, 能够用学习到的因果变量所对应的表征来构建对数据的表示, 并且通过干预和反事实操作来操纵数据, 再根据数据的变化来解释被操纵的隐变量的物理意义; 同时, 有限的因果变量也能够被对齐到真实世界中的重要概念, 帮助对表征的意义进行解释。

(3) **提升地理空间网络表征的稳定性。**因果机制内生地具有稳定不变的性质^[83], 这是因果关系与其他关系的区别所在。因此, 如果能够建模数据背后的因果机制, 那么这种因果关系可以作为推理链的一部分, 排除掉其余非因果变量的干扰, 根据已知的因果变量的信息推理出未知的因果变量的分布。这种稳定性为人流预测等地理空间网络的应用任务提供了基础, 如 STNSCM^[84]利用历史观测数据中捕获到的自行车流数据后隐含的因果机制, 结合未来时刻的时空上下文分布预测下一时刻的自行车流量。

(4) **提升地理空间网络表征的泛化性。**正如同人对于新环境的快速适应得益于熟知一系列关于物理世界的不变性规律, 地理空间网络表征模型的泛化能力也能够由学习到不变的结构知识来保证。由于因果知识描述的是数据生成背后最基本的作用机制, 因而能够表示最本质的概念之间的最简关系, 当数据的分布改变时, 只需要改变对其中少部分变量的设定, 其余的变量会随着因果关系链中传递的因果效应而自然地发生改变。因此, 这是一种能够较小的风险泛化到新环境的有效方式, 如 STEVE^[85]在实验中发现因果表征能够更好地泛化到时间分布外泛化和空间分布外泛化的数据上。

3.2 空间因果表征学习

空间因果表征学习的核心在于对地理实体间的因果关系进行建模, 并基于这种因果关系构建出不同地理分布的地理实体的特征。例如, 如何通过一部分地理实体的特征, 更好地预测另一部分地理实体的特征。本节介绍了从空间因果的视角进行地理空间网络表征的模型, 重点主要在于如何对地理实体进行因果发现, 并建模到表征过程中。现有的方法主要可以归结为两类: (1) 假设驱动的地理实体因果发现方法: 根据对于因果机制或性质的假设, 利用已有的因果发现方法, 从地理实体观测特征中探测出因果关系; (2) 数据驱动的地理实体因果发现方法: 假设地理实体间存在严格的因果关系, 这种关系可以指示空间上的物质、信息流动或其他影响关系, 并且能够以数据驱动的方式学习。

(1) **假设驱动的地理实体因果发现方法。**这类方法以因果机制或性质假设作为驱动, 并利用基于此假设的因果发现方法, 识别出比时空相关性更严格的因果性关系。SCDGCN^[86]利用 PCMCI^[87]方法去除了工业传感器网络中由共同的驱动传感器或间接途径导致的伪关联, 得到了一个更稀疏的因果图结构; 转移熵可以刻画变量之间信息的传递过程和方向, 被认为是识别因果关系的有效方式之一, 如二次修正转移熵被用于识别机场延误网络中的延误因果图, 因果图可以直接输入到时态 GCN 中进行时空特征的提取^[88]; 变量-滞后转移熵也被用于信号基站之间的因果关系识别, DIVC-GAT^[89]将每个信号基站的通信流量分解为趋势、周期和随机三个成分, 并针对每个成分学习因果结构; PSTN^[90]在转移熵的基础上提出了一个最优因果熵原则, 用于学习交通节点之间的因果网络结构; GT-CausIn^[91]利用了迭代因果发现算法^[92]; ST-CGI^[93]和 STGC-GNNs^[78]都使用格兰杰因果检验方法的思想作为空间因果发现的基本原则, 不同的是 ST-CGI 使用一个端到端的神经网络同时实现格兰杰因果发现和自回归预测; STGC-GNNs 则通过改良的时空格兰杰因果检验方法学习因果图, 并可以将其融合到不同的时空图神经预测模型上; CASPER^[94]则在格兰杰因果检验的基础上适应一个时空因果注意力模块对因果关系进一步稀疏化; 为了构建异常之间的因果关系, 帮助实现异常事件的探测^[95], 空间连续性、时间连续性和时间先后性等性质约束被定义用于探测因果关系。假设驱动的地理实体因果发现方法能够基于对因果机制的假设识别出满足假设条件的因果知识, 识别出的关系具有明确的性质, 理论性非常明确。然而, 这类方法在当前数据不符合该假设的情况下可能会失效, 如对线性数据中因果关系的假设并不一定会在非线性数据中满足。

(2) **数据驱动的地理实体因果发现方法。**这类方法以模型效能为驱动, 用于提取比时空相关性更严格的关联关系。STREAMS^[96]利用强化学习来学习因果图结构, 利用一个强化学习智能体逐步优化因果关系的搜索空间; STCTN^[97], STCGAT^[98]和 CSTGNN^[99]利用一个自学习的因果传递矩阵并实现端到端的优化; 与之类似的是 STNSCM^[84]使用了一个可学习的动态因果生成器, 可以根据时间序列观测值和当前外部环境, 生成区域间时变性的因果关系, 自注意力机制也被用于学习这种动态空间因果关系^[100]; CityCAN^[101]结合一个干预模块分别学习具有不变性的因果邻接矩阵和另一个无用矩阵, 并利用不变关系矩阵进行预测, CISTGNN^[102]也用类似的解耦策略, 利用空间注意力机制分别学习因果矩阵和捷径矩阵; 而 CauSTG^[103]认为在不同时态子环境下因果关系应当具有不变性, 因此设计了一种层次不变性探索器来发现不同子环境下模型的稳定权重, 实现对不变因果关系的发现, 该方法提供了在交通路网、环境污染网络和用电网络上的应用; CaST^[104]利用一个线性变换层从实体表征中学习不同地理实体间的因果强度, 并进一步利用了前门

调整去除空间聚合过程中空间混杂变量造成的伪相关；CGT^[105]则将交通路网中由多个交通传感器组成的子图视为原因变量，并利用扰动掩膜得到反事实图，再通过反事实解释器找到最重要的子图作为最终的原因变量。数据驱动的地理实体因果发现方法具有更强的数据适应性，能够以自适应的方式捕获到更逼近学习目标的因果关系，然而，这类方法相比于假设驱动的方法理论支撑较弱，且当观测数据中有误差或偏差时具有较高的风险。

3.3 时间因果表征学习

时间因果表征学习的核心在于对隐变量间的因果关系进行建模，并通过这种因果关系构建模型学习过程中隐变量的变换。例如，如何通过一部分隐变量的信息构建出另一部分隐变量的取值分布。本节介绍了从时间因果的视角进行地理空间网络表征的模型，重点在于如何刻画不同的隐变量及其之间的因果关系，并建模到表征过程中。根据针对隐变量的因果学习任务的不同，现有的方法主要可以归结为两类：(1) 基于隐变量因果发现的表征方法：通过学习隐变量间的因果关系构建对关注的时空变量的表征并实现下游任务；(2) 基于隐变量因果推断的表征方法：通过除混杂因子的影响推断隐变量的因果效应，得到地理空间网络的无混杂表征。

(1) 基于隐变量因果发现的表征方法。这类方法通过学习隐变量间的因果关系构建对关注的时空变量的预测。如 CausalGNN^[106]将每个地区的疑似、感染、康复和死亡病例数这四个状态特征直接作为想要学习的抽象隐变量，用一个因果编码器学习四个特征的嵌入并使用一个面向预测任务的因果解码器学习四个状态之间的转移概率作为因果结构；GCIM^[107]假设观测对象是由微观的因果描述子经过复杂的非线性时空混合过程生成的，为了学习出这些因果描述子及其之间的关系，可逆神经样条流 (invertible neural spline flows) 被用于拟合因果传递函数，并将学习过程直接融合到一个基于 GraphGRU^[108]的端到端预测框架中；CCHMM^[109]则首先预定了一个时态因果图骨架，然后再基于观测数据和外部信息对其中的因果邻接关系进行学习。

(2) 基于隐变量因果推断的表征方法。这类方法通过推断隐变量的因果效应去除混杂因子在时空预测中的影响，即主要关注实现因果效应估计这一子任务的无混杂性。不同的工作处理的混杂因子不同，去混杂的方法也不尽相同，根据所使用的因果推断框架的不同，这类方式可以进一步分为基于结构因果模型和潜在结果框架的方法。结构因果模型和潜在结果框架可以视为两种不同的建模语言，结构因果模型利用图形化的表示直观绘制因果问题并基于图论实现因果性操作；而潜在结果模型则利用公式化的计算语言基于统一的假设体系来获得因果性结论。两者各有所长且背后的因果推断理念并不冲突，研究者可以根据问题场景灵活选择。

基于结构因果模型的方法。这类方法能够通过直接在因果图结构上实施干预算子，将待建模的因果问题通过前门调整、后门调整等方式转换为能对数据进行表达的统计语言。后门调整通过对混杂因子进行分层和平衡实现去混杂，被诸多工作采纳。部分工作将混杂因子分类至预设的层数，如 STCTN^[97]认为区域人流网络中的区域属性是历史数据和未来数据的混杂因子，会使得现有时空预测模型在提取时域特征时吸收历史数据和未来数据间的虚假相关性。因此，使用了谱聚类对区域属性进行分层，并在独立参数空间中为每一层构建了特有的无偏预测模型用于实现无混杂的预测；SEAD^[110]的目标是去除社会环境对行人历史轨迹和未来轨迹造成的混杂效应，与之不同的是，混杂因子被假设服从均匀分布，且混杂因子与原因变量的联合分布通过一个交叉注意力模块来建；CISTGNN^[102]从外部天气环境中学习混杂因子的表示并通过后门准则去混杂；STEVE^[85]重点考虑了混杂因子层数的有限性和类别的完备性，将混杂因子分为不变性和变化性两层；CaST^[104]则通过一个时态环境编码本对混杂因子进行离散化，使得混杂表征能够落入最匹配的层中；CTSGI^[109]利用行人轨迹处的图片作为混杂因子表征的来源，使用语义分割模型从中提取该轨迹点的环境表征作为混杂，再使用后门准则对混杂效应进行调整。前门调整则适用了后门路径不存在的因果图中，提供了与后门准则不同的去混杂策略。STNSCM^[84]利用门控线性单元从历史观测数据和外部环境数据中得到干预后原因变量的分布，在该分布中，原因变量能够以无偏的概率与不同的混杂因子组合，从而实现去混杂的效应估计；类似的策略也被用于车辆轨迹的预测，其中场景上下文作为混杂经过一个基于前门调整策略的反事实表征推断模块在预测中被消除^[100]；CASPER^[94]利用前门调整将时间序列补全任务转变为对输入数据表征子类别的求和，并通过可学习的提示向量计算去混杂的因果效应。

基于潜在结果框架的方法。这类方法的主要理念则是将不同的观测单元视为样本，通过对样本的综合处理实现对干预操作的模拟，再进一步得到无混杂的因果效应。样本重加权是实现样本平衡化的经典方式之一^[111]，被用于将不同城市和地区分配到不同的处理组并计算不同处理对应的因果效应，如研究 POI 的变化对区域人流的影响^[112]。随着表征学习的发展，利用表征学习得到样本的统一表征后再进行混杂因子的平衡，也能够保证不同实验组的样本的背景特征分布相近。例如，CAPE^[113]为不同地点从历史事件序列中生成表征，再通过积分概率度量 (Integral Probability Metric, IPM) 度量表征上的差异，使得不同实验组的样本整体差异最小来实现表征平衡；SINet^[114]则利用希尔伯特-施密特独立性指标 (Hilbert-Schmidt

Independence Criterion, HSIC) 作为正则项引导表征平衡, 对隐藏的混杂因子全都进行去除; CIDER^[115]则利用 Wasserstein-1 距离对不同行政区的表征进行平衡。

4 挑战与展望

(1) **观测手段的局限性。**大部分地理空间网络的本质是时空数据, 在数据采样过程中需要针对所关注的特征沿时间轴的持续性观测和沿空间轴的大范围采样。然而, 由于外部环境的限制, 完整的时空特征的观测难度非常大, 且对于同一个观测对象而言, 一种观测手段往往只能刻画一类时空特征。如, 街景图像刻画城市的视觉特征, 而 POI 则刻画城市的功能分布, 新闻数据则描述了城市的动态活动。多模态学习为整合多源异构观测数据信息提供了契机, 未来, 如何利用多模态学习, 结合多源异构时空大数据^[116]揭示地理空间网络更全面、精准的机理, 是值得探索的重要方向。

(2) **地理过程的复杂性。**地理空间网络数据属于地理数据的一种, 理解复杂地理现象背后的机理^[117]是地理学研究的重要任务。然而, 地理过程并非单一机制作驱动下的结果, 而是地球内动力、地球外动力和人类活动驱动力复杂叠加下的结果。虽然因果学习具有揭示复杂、动态系统背后机理的潜能, 但针对复杂和动态演化的时空地理数据^[118], 实现对数据生成背后的多机制过程的解缠和可解释性建模仍然具有挑战。社会科学领域关于归因理论^[119]的研究旨在寻找行为和事件背后的原因, 本质上也是用因果的观点解释社会行为, 在社交网络等数据上已有成熟的研究, 能够帮助理解和解释地理空间网络数据中人类活动。未来如果加入对地球物理过程的先验知识和人地交互过程的专家理解作为对因果学习的约束, 或许能够帮助发现和解释更复杂的地理空间网络因果。

(3) **空间效应的特殊性。**地理空间分析中存在着空间异质性效应、空间交互的距离衰减效应、空间依赖近邻效应和空间分区的尺度效应这四种空间效应^[120]。然而, 因果学习领域认同的许多基本假设^[111]在空间效应存在时需要谨慎地斟酌^[121]。例如, 空间异质性效应会打破可忽略性假设, 空间分区的尺度效应可能会使得因果不变性存在可验证的边界, 而空间交互和空间依赖会导致空间溢出从而打破稳定单元干预值 (Stable Unit Treatment Value Assumption, SUTVA) 假设^[122], 这一问题在地理空间网络数据中会更加显著。如何在现有因果分析理论的基础上, 考虑地理空间网络数据的特性并建模, 如建模空间溢出效应^[123]和空间潜在混杂^[124, 125]等, 是未来地理空间网络因果表征学习中的重要问题。

5 结语

因果学习能够从数据中发现因果知识, 表征学习能够进一步对因果知识进行表达, 因果视角下的地理空间网络表征学习对于地理空间网络的研究与应用具有重要意义, 如关系挖掘和知识发现、提升下游模型的泛化性和可解释性等。本综述从因果视角对地理空间网络的表征学习进行了重新审视, 基于对现有地理空间网络表征学习框架的统一抽象, 提炼出了地理实体和隐变量两个不同层级的概念作为因果研究的对象, 并进一步演绎出了地理空间网络因果的两种表现形式——空间因果和时间因果。为了保证因果学习视角的系统性与统一性, 本文以结构因果模型为主线, 对地理空间网络因果学习任务进行了形式化的分解, 并基于该框架对现有工作进行了组织与分析。地理空间网络上的因果表征仍处于初步探索阶段, 也面临着本领域特有的挑战。本综述进一步对现存挑战和潜在的解决方式进行了探讨与展望, 未来期待着更多学者对地理空间网络因果表征学习的基本和重要的问题进行研究与思考。

表 3 地理空间网络空间因果表征学习已有研究

Tab.3 Existing works related to Geospatial Network Spatial Causality Representation Learning

文献	地理空间网络表征 (第 1.3 节)			地理空间网络因果学习 (第 2.3 节)		
	数据	表征模型	下游任务	因果变量表征	因果结构发现	因果效应估计
[86]	工业生产传感器网络	门控 TCN + GCN	传感器检测值预测	传感器时间序列观测值	PCMCI	—
[96]	水文网络	基于 LSTM+GCN 的时空编码器-解码器结构	流量预测	LSTM + 时态注意力 + GCN	基于强化学习的因果发现	—
[88]	机场延误传播网络	GRU+GCN	飞机延误预测	航班延误时间序列观测值	二次修正转移熵	—
[89]	通信基站网络	GRU+GAT	蜂窝通信流量预测	基站通信流量时间序列观测值分解后的趋势、周期和随机三个成分	变量-滞后转移熵	—
[90]	交通路网	基于一维 CNN + GRU + GCN 的时空表征块	交通流预测	交通流时间序列观测值	最优因果熵原则	—
[91]	交通路网	扩张因果卷积+图扩散卷积	交通流预测	每个时间窗内的车速变化	迭代因果发现	—
[93]	交通路网	扩张卷积 + 单层 GCN	交通流预测	扩张因果卷积	格兰杰因果检验	—
[78]	交通路网	基于 GNN 的交通流预测模型	交通流预测	交通流时间序列观测值+时空对齐算法	时空格兰杰因果检验	—
[94]	PM2.5 空气污染网络, 交通路网	Transformer+时空注意力机制	时间序列补全	时态 Transformer 编码器	格兰杰因果检验+时空因果注意力	—
[97]	区域人流网络	GRU+因果传递网络	出租车需求量预测	人流序列观测值	基于反向传播的自适应矩阵学习	—
[98]	交通路网	节点自适应学习+ GAT + 时态卷积	交通流预测	节点自适应学习 + GAT	基于双向 GRU 的可学习的图结构	—
[102]	交通路网	时态注意力 + 空间注意力 + GNN	交通流预测	时态注意力	解耦空间注意力机制	—

[99]	锅炉传感器网络	时态卷积 + 多跳 GCN	温度预测	温度时间序列观测值	可学习的结构因果方程	—
[84]	区域人流网络	时空进化 GCN	人流预测	外部环境和观察值的拼接	可学习的动态因果生成器	—
[100]	交通路网	LSTM + GCN + 自注意力 + GRU	车辆轨迹预测	历史轨迹和场景的融合表征	自注意力机制	—
[101]	区域交通网络	全局和局部自注意力编码器	时空预测	二维 CNN	干预模块+可学习的不变因果邻接图	—
[103]	交通路网, PM2.5 环境污染网络, 用电网络	时序分解+ TCN + GCN	时空预测	每个子环境下分别建立的回归模型	层次不变性探索器	—
[95]	区域网络	异常树	异常事件探测	基于区域与交通轨迹提取的异常	基于规则的因果发现	—
[105]	交通路网	时空图 Transformer	交通流预测	扰动掩膜	反事实解释器	—
[104]	交通路网, PM2.5 环境污染网络	TCN + GCN	时空预测	从时间序列观测值中解耦出来的实体表征	线性变换层	前门准则

表 4 地理空间网络时间因果表征学习已有研究

Tab.4 Existing works related to Geospatial Network Temporal Causality Representation Learning

文献	地理空间网络表征 (第 1.3 节)			地理空间网络因果学习 (第 2.3 节)		
	数据	表征模型	下游任务	因果变量表征	因果结构发现	因果效应估计
[106]	疫情传播网络	基于注意力的动态 GNN	疫情人数预测	疑似、感染、康复和死亡人数构成的矩阵	面向预测的因果解码器	—
[107]	区域人流网络	GraphGRU	人流预测	内生变量: GraphGRU 表征观测数据; 外生变量: 多层感知机+ softmax 得到域类别;	可逆神经样条流拟合因果传递函数	—
[109]	区域人流网络	GraphGRU	人流预测	内生变量: 由外生因果变量和因果邻接矩阵建模; 外生变量: GraphGRU 表征观测数据和外部信息;	预定义的时态因果图骨架+可训练的因果邻接图	—
[97]	区域人流网络	GRU+因果传递网络	出租车需求量预测	混杂变量: 谱聚类得到不同区域属性的类簇; 原因和结果变量: 时空因果表示编码器;	—	后门调整
[110]	行人社交网络	不同的 STGNN, ST-Transformer	未来轨迹预测	混杂变量: 可学习的社会环境集; 原因和结果变量: 时空表征骨架模型;	—	后门调整
[85]	区域人流网络	一维因果卷积 + GCN	人流量预测	混杂变量: 历史数据的时空表征解耦; 原因和结果变量: 时空卷积层;	—	后门调整

[104]	交通路网, PM2.5 环境污染网络	TCN + GCN	时空预测	混杂变量: 历史数据的时空表征解耦; 原因和结果变量: 基于 TCN 的骨干编码器;	—	后门调整
[102]	交通路网	时态注意力+空间注意力+GNN	交通流预测	混杂变量: 天气条件; 原因变量: 历史交通流; 结果变量: 未来交通流;	—	后门调整
[84]	区域人流网络	时空进化 GCN	人流预测	混杂变量: 外部环境数据; 原因和结果变量: 门控线性单元; 混杂变量: GCN 表征场景上下文;	—	前门调整
[100]	交通路网	LSTM + GCN + 自注意力 + GRU	车辆轨迹预测	原因变量: 历史轨迹; 结果变量: 未来轨迹;	—	前门调整
[94]	PM2.5 空气污染网络, 交通路网	Transformer+时空注意力机制	时间序列补全	混杂变量: 可学习的提示变量; 原因变量: 输入的观测数据 结果变量: 待补全的观测数据;	—	前门调整
[112]	区域人流网络	OD 预测骨干模型 (重力模型, 图学习模型等)	OD 预测	混杂变量: 属性表征中解耦; 原因变量: POI 变化 + 变分自编码器; 结果变量: 未来人流;	—	变分样本重加权
[113]	社会事件网络	门控空洞因果卷积 + GCN	事件预测	混杂变量: 历史事件序列矩阵的表征; 原因变量: 改变过去发生的某类事件; 结果变量: 未来时间是否发生;	—	基于积分概率度量的表征平衡
[114]	层次空间图	Graph wavenet	导航成功率预测	混杂变量: GCN 提取场所表征; 原因变量: 路径导航规划; 结果变量: 到达目的地的智能体数量;	—	基于希尔伯特-施密特独立性指标正则项的表征平衡
[115]	传染病传播网络	RNN + GCN	疾病爆发预测	混杂变量: GCN 提取行政区表征; 原因变量: 疫情相关政策; 结果变量: 确诊和死亡病例数;	—	基于 Wasserstein-1 距离正则项的表征平衡
[109]	行人交互网络	多头注意力 + GCN	行人轨迹预测	混杂变量: 语义分割模型得到标签+卷积+池化; 原因变量: 历史轨迹; 结果变量: 未来轨迹;	—	后门调整

参考文献

- [1] TOBLER W R. A computer movie simulating urban growth in the Detroit region [J]. *Economic geography*, 1970, 46(sup1): 234-40.
- [2] 宋长青, 张国友, 程昌秀, 等. 论地理学的特性与基本问题 [J]. *地理科学*, 2020, 40(1): 6-11.
Song Changqing, Zhang Guoyou, Cheng Changxiu, et al. Nature and basic issues of Geography. *Scientia Geographica Sinica*, 2020, 40(1): 6-11.
- [3] HAINING R. Geography and spatial statistics: current positions, future developments [J]. *Remodelling geography*, 1989: 191-203.
- [4] CURTIN K M. Network Analysis in Geographic Information Science: Review, Assessment, and Projections [J]. *Cartography and Geographic Information Science*, 2013, 34: 9.
- [5] SCHÖTTLER S, YANG Y, PFISTER H, et al. Visualizing and interacting with geospatial networks: A survey and design space; proceedings of the Computer Graphics Forum, F, 2021 [C]. Wiley Online Library.
- [6] 刘彦随. 现代人地关系与人地系统科学 [J]. *地理科学*, 2020, 40(8): 1221-34.
Liu Yansui. Modern human-earth relationship and human-earth system science. *Scientia Geographica Sinica*, 2020, 40(8): 1221-1234.
- [7] 秦昆, 周勃, 徐源泉, 等. 城市交通热点区域的空间交互网络分析 [J]. *地理科学进展*, 2017, 36(9): 1149-1157.
QIN Kun, ZHOU Qing, XU Yuanquan, et al. Spatial interaction network analysis of urban traffic hotspots [J]. *Progress in Geography*, 2017, 36(9): 1149-1157.
- [8] 程昌秀, 史培军, 宋长青, 等. 地理大数据为地理复杂性研究提供新机遇 [J]. *地理学报*, 2018, 73(8): 1397-406.
CHENG Changxiu, SHI Peijun, SONG Changqing, et al. Geographic big-data: A new opportunity for geography complexity study [J]. *Acta Geographica Sinica*, 2018, 73(8): 1397-1406.
- [9] 刘瑜, 姚欣, 龚咏喜, 等. 大数据时代的空间交互分析方法和应用再论 [J]. *地理学报*, 2020, 75(7): 1523-1538.
LIU Yu, YAO Xin, GONG Yongxi, et al. Analytical methods and applications of spatial interactions in the era of big data [J]. *Acta Geographica Sinica*, 2020, 75(7): 1523-1538.
- [10] 刘瑜, 肖昱, 高松, 等. 基于位置感知设备的人类移动研究综述 [J]. *地理与地理信息科学*, 2011, 27(4): 8-13+31+2.
LIU Yu, XIAO Yu, GAO Song, et al. A Review of Human Mobility Research Based on Location Aware Devices [J]. *Geography and Geo-information Science*, 2011, 27(4): 8-13+31+2.
- [11] 杨学习, 邓敏, 刘瑜. 社会感知与地理空间智能的研究动态与展望——“社会感知与地理空间智能”专栏导读 [J]. *地理与地理信息科学*, 2022, 38(1).
YANG Xuexi, DENG Min, LIU Yu. Recent Advances and Future Directions of Social Sensing and GeoAI: Reading Guidance for the Special Columns of “Social Sensing and GeoAI” [J]. *Geography and Geo-information Science*, 2022, 38(1).
- [12] 高松. 地理空间人工智能的近期研究总结与思考 [J]. *武汉大学学报(信息科学版)*, 2020, 45(12).
GAO Song. A Review of Recent Researches and Reflections on Geospatial Artificial Intelligence [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(12): 1865-1874.
- [13] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: A review and new perspectives [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(8): 1798-828.
- [14] 吴俊杰, 刘冠男, 王静远, 等. 数据智能: 趋势与挑战. *系统工程理论与实践*, 2020, 40(8): 2116-2149
WU Junjie, LIU Guannan, WANG Jingyuan, ZUO Yuan, BU Hui, LIN Hao. Data intelligence: Trends and challenges. *Systems Engineering - Theory & Practice*, 2020, 40(8): 2116-2149
- [15] KIPF T N, WELLMING M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv:160902907, 2016.
- [16] RAHIMI A, COHN T, BALDWIN T. Semi-supervised User Geolocation via Graph Convolutional Networks; proceedings of the Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), F, 2018 [C].
- [17] ZHU D, ZHANG F, WANG S, et al. Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks [J]. *Annals of the American Association of Geographers*, 2020, 110(2): 408-20.
- [18] PEARL J, MACKENZIE D. *The Book of Why: The New Science of Cause and Effect* [M]. Basic Books, Inc., 2018.
- [19] WANG J, JIANG J, JIANG W, et al. LibCity: An Open Library for Traffic Prediction [Z]. Proceedings of the 29th International Conference on Advances in Geographic Information Systems. 2021.10.1145/3474717.3483923
- [20] CHOW C-Y, BAO J, MOKBEL M F. Towards location-based social networking services; proceedings of the proceedings of the 2nd ACM SIGSPATIAL International Workshop on location based social networks, F, 2010 [C].
- [21] HAYASHI Y. A review of recent studies of geographical scale-free networks [J]. *Information and Media Technologies*, 2006, 1(2): 1136-45.
- [22] GUIMERA R, MOSSA S, TURTSCHI A, et al. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles [J]. *Proceedings of the National Academy of Sciences*, 2005, 102(22): 7794-9.
- [23] IANNELLI F, KOHER A, BROCKMANN D, et al. Effective distances for epidemics spreading on complex networks [J]. *Physical Review E*, 2017, 95(1): 012313.
- [24] LIN Y, MAGO N, GAO Y, et al. Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning [Z]. Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Seattle, Washington. 2018.10.1145/3274895.3274907
- [25] ZHANG X, LI Y, FRERY A C, et al. Sea Surface Temperature Prediction With Memory Graph Convolutional Networks [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5.
- [26] 杨文龙, 杜德斌, 马亚华, 等. “一带一路”沿线国家贸易网络空间结构与邻近性 [J]. *地理研究*, 2018, 37(11): 2218-35.
YANG Wenlong, DU Debin, MA Yahua, et al. Network structure and proximity of the trade network in the Belt and Road region [J]. *Geographical Research*, 2018, 37(11): 2218-2235.
- [27] 吕峥, 孙群, 马京振, 等. 复杂网络视角下的居民地选取方法 [J]. *测绘学报*, 2023, 52(5): 852-62.
Lü Zheng, SUN Qun, MA Jingzhen, et al. Residential area selection method from the perspective of complex network [J]. *Acta Geodaetica et Cartographica Sinica*, 2023, 52(5): 852-862.
- [28] 秦昆, 喻雪松, 周扬, 等. 全球尺度地理多元流的网络化挖掘及关联分析研究 [J]. *地球信息科学学报*, 2022, 24(10): 1911-1924.
QIN Kun, YU Xuesong, ZHOU Yang, et al. Networked Mining and Association Analysis of Geographical Multiple Flows at a Global Scale [J]. *Geo-Information Science*, 2022, 24(10): 1911-1924.

-
- [29] BONACCORSI G, RICCABONI M, FAGIOLO G, et al. Country centrality in the international multiplex network [J]. *Applied Network Science*, 2019, 4: 1-42.
- [30] 梁天祺, 秦昆, 阮建平, 等. 基于多层复杂网络的地理多元流测度与社区识别研究[J]. *地球信息科学学报*, 2024, 26(8): 1843-1857.
LIANG Tianqi, QIN Kun, RUAN Jianping, et al. Research on Measurement and Community Detection of Geographic Multiple Flow Based on Multi-layer Network Methods[J]. *Journal of Geo-information Science*, 2024, 26(8): 1843-1857.
- [31] 王梓又. 基于多源异构数据流的成渝地区双城经济圈空间网络结构研究 [D], 2023.
WANG Ziyou. Spatial network structure of Chengdu-Chongqing Economic Circle based on Multi-source Heterogeneous Data Stream [D], 2023.
- [32] ALETA A, MELONI S, MORENO Y. A multilayer perspective for the analysis of urban transportation systems [J]. *Scientific reports*, 2017, 7(1): 44359.
- [33] 赵金丽, 张学波, 任嘉敏, 等. 多元流视角下黄河流域城市网络空间结构及其影响因素 [J]. *地理科学*, 2022, 42(10): 1778-1787.
ZHAO Jinli, ZHANG Xuebo, REN Jiamin, et al. Spatial structure and influencing factors of urban network in the Yellow River Basin based on multiple flows. *Scientia Geographica Sinica*, 2022, 42(10):1778-1787.
- [34] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk: Online learning of social representations; proceedings of the Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, F, 2014 [C].
- [35] GROVER A, LESKOVEC J. node2vec: Scalable feature learning for networks; proceedings of the Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, F, 2016 [C].
- [36] ZHU D, CUI P, ZHANG Z, et al. High-order proximity preserved embedding for dynamic networks [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2018, 30(11): 2134-44.
- [37] WANG X, CUI P, WANG J, et al. Community preserving network embedding; proceedings of the Thirty-first AAAI conference on artificial intelligence, F, 2017 [C].
- [38] TANG L, LIU H. Relational learning via latent social dimensions; proceedings of the Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, F, 2009 [C].
- [39] TANG L, LIU H. Leveraging social media networks for classification [J]. *Data Mining and Knowledge Discovery*, 2011, 23(3): 447-78.
- [40] TANG L, LIU H. Scalable learning of collective behavior based on sparse social dimensions; proceedings of the Proceedings of the 18th ACM conference on Information and knowledge management, F, 2009 [C].
- [41] RIBEIRO L F, SAVERESE P H, FIGUEIREDO D R. struc2vec: Learning node representations from structural identity; proceedings of the Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, F, 2017 [C].
- [42] DONNAT C, ZITNIK M, HALLAC D, et al. Spectral graph wavelets for structural role similarity in networks [J]. 2018.
- [43] FENG R, YANG Y, HU W, et al. Representation learning for scale-free networks; proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, F, 2018 [C].
- [44] CHEN H, PEROZZI B, HU Y, et al. Harp: Hierarchical representation learning for networks; proceedings of the Proceedings of the AAAI conference on artificial intelligence, F, 2018 [C].
- [45] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering [J]. *Advances in neural information processing systems*, 2016, 29.
- [46] LI R, WANG S, ZHU F, HUANG J. Adaptive graph convolutional neural networks; proceedings of the Proceedings of the AAAI conference on artificial intelligence, F, 2018 [C].
- [47] ZHUANG C, MA Q. Dual graph convolutional networks for graph-based semi-supervised classification; proceedings of the Proceedings of the 2018 world wide web conference, F, 2018 [C].
- [48] JOAN B, WOJCIECH Z, ARTHUR S, et al. Spectral networks and locally connected networks on graphs; proceedings of the International Conference on Learning Representations (ICLR2014), F, 2014 [C].
- [49] ATWOOD J, TOWSLEY D. Diffusion-convolutional neural networks [J]. *Advances in neural information processing systems*, 2016, 29.
- [50] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs [J]. *Advances in neural information processing systems*, 2017, 30.
- [51] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. *stat*, 2017, 1050(20): 10-48550.
- [52] DWIVEDI V P, BRESSON X. A generalization of transformer networks to graphs; proceedings of the Proceedings of the AAAI Workshop on Deep Learning on Graphs (AAAI 2021), F, 2021 [C].
- [53] KREUZER D, BEAINI D, HAMILTON W, et al. Rethinking graph transformers with spectral attention [J]. *Advances in Neural Information Processing Systems*, 2021, 34: 21618-29.
- [54] PARK J, YUN S, PARK H, et al. Deformable graph transformer [J]. *arXiv preprint arXiv:220614337*, 2022.
- [55] YING C, CAI T, LUO S, et al. Do transformers really perform badly for graph representation? [J]. *Advances in neural information processing systems*, 2021, 34: 28877-88.
- [56] MIALON G, CHEN D, SELOSSE M, et al. Graphit: Encoding graph structure in transformers [J]. *arXiv preprint arXiv:210605667*, 2021.
- [57] KONG K, CHEN J, KIRCHENBAUER J, et al. GOAT: A global transformer on large-scale graphs; proceedings of the International Conference on Machine Learning, F, 2023 [C]. PMLR.
- [58] MA L, LIN C, LIM D, et al. Graph inductive biases in transformers without message passing; proceedings of the International Conference on Machine Learning, F, 2023 [C]. PMLR.
- [59] CHAI Z, ZHANG T, WU L, et al. Graphllm: Boosting graph reasoning ability of large language model [J]. *arXiv preprint arXiv:231005845*, 2023.
- [60] FATEMI B, HALCROW J, PEROZZI B. Talk like a graph: Encoding graphs for large language models [J]. *arXiv preprint arXiv:231004560*, 2023.
- [61] GUO J, DU L, LIU H, et al. Gpt4graph: Can large language models understand graph structured data? an empirical evaluation and benchmarking [J]. *arXiv preprint arXiv:230515066*, 2023.
- [62] LIU C, WU B. Evaluating large language models on graphs: Performance insights and comparative analysis [J]. *arXiv preprint arXiv:230811224*, 2023.
- [63] WANG H, FENG S, HE T, et al. Can language models solve graph problems in natural language? [J]. *Advances in Neural Information Processing Systems*, 2024, 36.
- [64] JIANG W, LUO J. Graph neural network for traffic forecasting: A survey [J]. *Expert Systems with Applications*, 2022, 207: 117921.

-
- [65] ZHAO S, LIU R, CHENG B, et al. Classification-Labelled Continuousization and Multi-Domain Spatio-Temporal Fusion for Fine-Grained Urban Crime Prediction [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(7): 6725-38.
- [66] VARSHNEY Y, KUMAR V, DUBEY D K, et al. Forecasting Precision: The Role of Graph Neural Networks and Dynamic GNNs in Weather Prediction [J]. *Journal of Big Data Technology and Business Analytics*, 2024, 3(1): 28-33.
- [67] MYLAVARAPU S, SANDHU M, VIJAYAN P, et al. Towards Accurate Vehicle Behaviour Classification With Multi-Relational Graph Convolutional Networks; proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), F 19 Oct.-13 Nov. 2020, 2020 [C].
- [68] 朱强. 基于时空大数据的城市人群流动预测研究 [D], 2020.
- ZHU Qiang. Study on the Citywide Crowd Flow Prediction based on Spatial Temporal Big Data [D], 2020.
- [69] YAO X, GAO Y, ZHU D, et al. Spatial Origin-Destination Flow Imputation Using Graph Convolutional Networks [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(12): 7474-84.
- [70] YUAN S, WANG C, JIANG Q, et al. Community Detection with Graph Neural Network using Markov Stability; proceedings of the 2022 International Conference on Artificial Intelligence in Information and Communication (ICAIC), F 21-24 Feb. 2022, 2022 [C].
- [71] ROMANOVA A. GNN Graph Classification Method to Discover Climate Change Patterns; proceedings of the Artificial Neural Networks and Machine Learning – ICANN 2023, Cham, F 2023//, 2023 [C]. Springer Nature Switzerland.
- [72] PEARL J, GLYMOUR M, JEWELL N P. Causal inference in statistics: A primer [M]. John Wiley & Sons, 2016.
- [73] PEARL J. Causality: Models, reasoning and inference [J]. Cambridge, UK: Cambridge University Press, 2000, 19(2): 3.
- [74] YANG Y, NAFAE M S, GHASSAMI A, et al. Causal Discovery in Linear Structural Causal Models with Deterministic Relations [Z]//BERNHARD S, CAROLINE U, KUN Z. Proceedings of the First Conference on Causal Learning and Reasoning. Proceedings of Machine Learning Research; PMLR. 2022: 944--93
- [75] 马忠贵, 徐晓晗, 刘雪儿. 因果推断三种分析框架及其应用综述 [J]. *工程科学学报*, 2022, 44(7): 1231-43.
- MA Zhonggui, XU Xiaohan, LIU Xueer. Three analytical frameworks of causal inference and their applications [J]. *Chinese Journal of Engineering*, 2022, 44(7): 1231-1243.
- [76] YANG M, LIU F, CHEN Z, et al. Causalvae: Disentangled representation learning via neural structural causal models; proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, F, 2021 [C].
- [77] CHOI S, JEONG M, HAN H, et al. C2L: Causally Contrastive Learning for Robust Text Classification [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(10): 10526-34.
- [78] HE S, LUO Q, DU R, et al. STGC-GNNs: A GNN-based traffic prediction framework with a spatial-temporal Granger causality graph [J]. *Physica A: Statistical Mechanics and its Applications*, 2023, 623: 128913.
- [79] RUNGE J, GERHARDUS A, VARANDO G, et al. Causal inference for time series [J]. *Nature Reviews Earth & Environment*, 2023, 4(7): 487-505.
- [80] 蔡瑞初, 陈薇, 张坤, 等. 基于非时序观察数据的因果关系发现综述 [J]. *计算机学报*, 2017, 40(6): 1470-90.
- CAI Ruichu, CHEN Wei, ZHANG Kun, et al. A Survey on Non-Temporal Series Observational Data Based Causal Discovery [J]. *Chinese Journal of Computers*, 2017, 40(6): 1470-90.
- [81] NOGUEIRA A R, PUGNANA A, RUGGIERI S, et al. Methods and tools for causal discovery and causal inference [J]. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 2022, 12(2): e1449.
- [82] 吴鑫然, 岳昆. 带隐变量贝叶斯网学习方法: 研究综述 [J]. *云南大学学报(自然科学版)*, 2023, 45(2): 298-313.
- WU Xin-ran, YUE Kun. Approaches for learning a Bayesian network with latent variables: A survey [J]. *Journal of Yunnan University: Natural Sciences Edition*, 2023, 45(2): 298-313.
- [83] SCHÖLKOPF B, LOCATELLO F, BAUER S, et al. Toward Causal Representation Learning [J]. *Proceedings of the IEEE*, 2021, 109(5): 612-34.
- [84] DENG P, ZHAO Y, LIU J, et al. Spatio-temporal neural structural causal models for bike flow prediction; proceedings of the Proceedings of the AAAI conference on artificial intelligence, F, 2023 [C].
- [85] JI J, ZHANG W, WANG J, et al. Self-Supervised Deconfounding Against Spatio-Temporal Shifts: Theory and Modeling [J]. *arXiv preprint arXiv:231112472*, 2023.
- [86] ZHENG S, HAO K, SHI X, et al. Spatiotemporal Causal Discovery Graph Convolutional Networks for Multivariate Time Series Forecasting of Industrial Process; proceedings of the 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), F, 2023 [C]. IEEE.
- [87] RUNGE J, BATHIANY S, BOLLE E, et al. Inferring causation from time series in Earth system sciences [J]. *Nature Communications*, 2019, 10(1): 2553.
- [88] SUN M, TIAN Y, WANG X, et al. Transport causality knowledge-guided GCN for propagated delay prediction in airport delay propagation networks [J]. *Expert Systems with Applications*, 2024, 240: 122426.
- [89] CHEN X, CHUAI G, ZHANG K, et al. Spatial-temporal cellular traffic prediction: A novel method based on causality and graph attention network; proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC), F, 2023 [C]. IEEE.
- [90] LIU X, SHAO Q, CHEN D. Long-term Prediction on Graph Data with Causal Network Construction [J]. *IEEE Transactions on Artificial Intelligence*, 2024.
- [91] GAO T, MARQUES R K, YU L. GT-CausIn: a novel causal-based insight for traffic prediction [J]. *arXiv preprint arXiv:221205782*, 2022.
- [92] ROHEKAR R Y, NISIMOV S, GURWICZ Y, et al. Iterative causal discovery in the possible presence of latent confounders and selection bias [J]. *Advances in Neural Information Processing Systems*, 2021, 34: 2454-65.
- [93] ZHANG L, FU K, JI T, et al. Granger causal inference for interpretable traffic prediction; proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), F, 2022 [C]. IEEE.
- [94] JING B, ZHOU D, REN K, et al. CASPER: Causality-Aware Spatiotemporal Graph Neural Networks for Spatiotemporal Time Series Imputation; proceedings of the the 33rd ACM International Conference on Information and Knowledge Management (CIKM ' 24), New York, US, F October 21 - 25, 2024, 2024 [C].
- [95] LIU W, ZHENG Y, CHAWLA S, et al. Discovering spatio-temporal causal interactions in traffic data streams; proceedings of the Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, F, 2011 [C].
- [96] SHETH P, MOSALLANEZHAD A, DING K, et al. STREAMS: Towards Spatio-Temporal Causal Discovery with Reinforcement Learning for Streamflow Rate Prediction; proceedings of the Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, F,

2023 [C].

[97] 邓攀, 刘俊廷, 王晓, 等. STCTN: 一种基于时域偏倚校正与空域因果传递的时空因果表示学习方法 [J]. 计算机学报, 2023, 46(12): 2535-50.

DENG Pan, LIU Junting, WANG Xiao, et al. STCTN: a Spatio-temporal Causal Representation Learning Method Based on Temporal Bias Adjustment and Spatial Causal Transition [J]. Chinese Journal of Computers, 2023, 46(12): 2535-50.

[98] ZHAO W, ZHANG S, WANG B, et al. Spatio-temporal causal graph attention network for traffic flow prediction in intelligent transportation systems [J]. PeerJ Computer Science, 2023, 9: e1484.

[99] MIRAKI A, DAPKUTĚ A, ŠIOŽINYS V, et al. Causal-Based Spatio-Temporal Graph Neural Networks for Industrial Internet of Things Multivariate Time Series Forecasting; proceedings of the Explainable Artificial Intelligence, Cham, F 2023//, 2023 [C]. Springer Nature Switzerland.

[100] LIU J, LIN H, WANG X, et al. Reliable trajectory prediction in scene fusion based on spatio-temporal Structure Causal Model [J]. Information Fusion, 2024, 107: 102309.

[101] WANG C, LIANG Y, TAN G. CityCAN: Causal Attention Network for Citywide Spatio-Temporal Forecasting [Z]. Proceedings of the 17th ACM International Conference on Web Search and Data Mining. Merida, Mexico; Association for Computing Machinery. 2024: 702–11.10.1145/3616855.3635764

[102] LUO X, YIN W, LI Z. Spatio-Temporal Graph Neural Network with Hidden Confounders for Causal Forecast; proceedings of the Special Track on AI for Socio-Ecological Welfare at ICCBR2024 (ICCBRAI Track'24), Merida, Mexico, F July 1, 2024, 2024 [C].

[103] ZHOU Z, HUANG Q, YANG K, et al. Maintaining the Status Quo: Capturing Invariant Relations for OOD Spatiotemporal Learning [Z]. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Long Beach, CA, USA; Association for Computing Machinery. 2023: 3603–14.10.1145/3580305.3599421

[104] XIA Y, LIANG Y, WEN H, et al. Deciphering spatio-temporal graph forecasting: A causal lens and treatment [J]. Advances in Neural Information Processing Systems, 2024, 36.

[105] YANG Y, DU K, DAI X, et al. Counterfactual Graph Transformer for Traffic Flow Prediction; proceedings of the 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), F, 2023 [C]. IEEE.

[106] WANG L, ADIGA A, CHEN J, et al. Causalgnn: Causal-based graph neural networks for spatio-temporal epidemic forecasting; proceedings of the Proceedings of the AAAI conference on artificial intelligence, F, 2022 [C].

[107] ZHAO Y, DENG P, LIU J, et al. Generative Causal Interpretation Model for Spatio-Temporal Representation Learning; proceedings of the Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, F, 2023 [C].

[108] HE H, SU L, YE K. GraphGRU: A Graph Neural Network Model for Resource Prediction in Microservice Cluster; proceedings of the 2022 IEEE 28th International Conference on Parallel and Distributed Systems (ICPADS), F 10-12 Jan. 2023, 2023 [C].

[109] ZHAO Y, DENG P, LIU J, et al. Causal conditional hidden Markov model for multimodal traffic prediction; proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, F, 2023 [C].

[110] GE C, SONG S, HUANG G. Causal intervention for human trajectory prediction with cross attention mechanism; proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, F, 2023 [C].

[111] YAO L, CHU Z, LI S, et al. A survey on causal inference [J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2021, 15(5): 1-46.

[112] ZENG J, ZHANG G, RONG C, et al. Causal learning empowered OD prediction for urban planning; proceedings of the Proceedings of the 31st ACM International Conference on Information & Knowledge Management, F, 2022 [C].

[113] DENG S, RANGWALA H, NING Y. Robust event forecasting with spatiotemporal confounder learning; proceedings of the Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, F, 2022 [C].

[114] TAKEUCHI K, NISHIDA R, KASHIMA H, et al. Causal effect estimation on hierarchical spatial graph data; proceedings of the Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, F, 2023 [C].

[115] MA J, DONG Y, HUANG Z, et al. Assessing the causal impact of COVID-19 related policies on outbreak dynamics: A case study in the US; proceedings of the Proceedings of the ACM Web Conference 2022, F, 2022 [C].

[116] 王家耀. 时空大数据时代的地图学 [J]. 测绘学报, 2017, 46(10): 1226-37.

WANG Jiayao. Cartography in the Age of Spatio-temporal Big Data[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(10): 1226-1237.

[117] 刘瑜, 郭浩, 李海峰, 等. 从地理规律到地理空间人工智能 [J]. 测绘学报, 2022, 51(6).

LIU Yu, GUO Hao, LI Haifeng, et al. A note on GeoAI from the perspective of geographical laws[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(6): 1062-1069.

[118] 李德仁, 马军, 邵振峰. 论时空大数据及其应用[J]. 卫星应用, 2015, 0(9):7-11.

LI Deren, MA Jun, SHAO Zhenfeng. Spatiotemporal Big Data and its Application[J]. Satellite Application. 2015, 0(9):7-11.

[119] DAVIS M L, WITCRAFT S M, BAIRD S O, et al. Chapter 3 - Learning Principles in CBT [M]/HOFMANN S G, ASMUNDSON G J G. The Science of Cognitive Behavioral Therapy. San Diego; Academic Press. 2017: 51-76.

[120] 刘瑜, 汪珂丽, 邢满月, 等. 地理分析中的空间效应[J]. 地理学报, 2023, 78(3): 517-531.

LIU Yu, WANG Keli, XING Xiaoyue, et al. On spatial effects in geographical analysis[J]. Acta Geographica Sinica, 2023, 78(3): 517-531

[121] GAO B, WANG J, STEIN A, et al. Causal inference in spatial statistics [J]. Spatial Statistics, 2022, 50: 100621.

[122] 李海峰, 罗琴瑶, 贺丝露, 等. 地理空间因果原则及地理空间效应的因果发现[J]. 武汉大学学报 (信息科学版), 2023, 48(11): 1800-1812.

LI Haifeng, LUO Qinyao, HE Silu, et al. Geospatial Causal Principle and Causal Discovery for Geospatial Effects[J]. Geomatics and Information Science of Wuhan University, 2023, 48(11): 1800-1812.

[123] REICH B J, YANG S, GUAN Y, et al. A Review of Spatial Causal Inference Methods for Environmental and Epidemiological Applications [J]. International Statistical Review, 2021, 89(3): 605-34.

[124] URDANGARIN A, GOICOA T, UGARTE M D. Evaluating recent methods to overcome spatial confounding [J]. Revista Matemática Complutense, 2023, 36(2): 333-60.

[125] URDANGARIN A, GOICOA T, KNEIB T, et al. A simplified spatial+ approach to mitigate spatial confounding in multivariate spatial areal models [J]. Spatial Statistics, 2024, 59: 100804.

网络首发:

标题: 因果视角下的地理空间网络表征方法综述: 进展、挑战与展望

作者: 李海峰, 贺丝露, 陈海鹏, 刘瑜, 顾鑫

收稿日期: 2024-12-22

DOI:10.13203/j.whugis20240366

引用格式:

李海峰, 贺丝露, 陈海鹏, 等. 因果视角下的地理空间网络表征方法综述: 进展、挑战与展望[J]. 武汉大学学报(信息科学版), 2025, DOI:10.13203/J.whugis20240366 (LI Haifeng, HE Silu, CHEN Haipeng, et al. Survey on Geospatial Network Representation Learning from a Causal Perspective: Advances, Challenges, and Prospects[J]. Geomatics and Information Science of Wuhan University, 2025, DOI:10.13203/J.whugis20240366)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

地理空间因果原则及地理空间效应的因果发现

李海峰, 罗琴瑶, 贺丝露, 任臻, 刘瑜

武汉大学学报(信息科学版), 2023, 48(11): 1800-1812.

<http://ch.whu.edu.cn/article/doi/10.13203/j.whugis20230351>

网络空间地图可视化方法研究综述

张兰, 王光霞, 蒋秉川, 张蓝天, 马麟

武汉大学学报(信息科学版), 2022, 47(12): 2113-2122.

<http://ch.whu.edu.cn/article/doi/10.13203/j.whugis20220244>