



(12) 发明专利申请

(10) 申请公布号 CN 121146017 A

(43) 申请公布日 2025. 12. 16

(21) 申请号 202511271490.3

G06F 16/29 (2019.01)

(22) 申请日 2025.09.08

(71) 申请人 广东省城乡规划设计研究院科技集团股份有限公司

地址 510290 广东省广州市海珠区南洲路483号

(72) 发明人 张经度 魏冀明 廖瑞毅 龚双
陈艳 张洪剑 连臣 陈得冠
王晖 周元

(74) 专利代理机构 广州粤高专利商标代理有限公司 44102

专利代理师 吴远鹏

(51) Int. Cl.

G06N 5/022 (2023.01)

G06N 5/02 (2023.01)

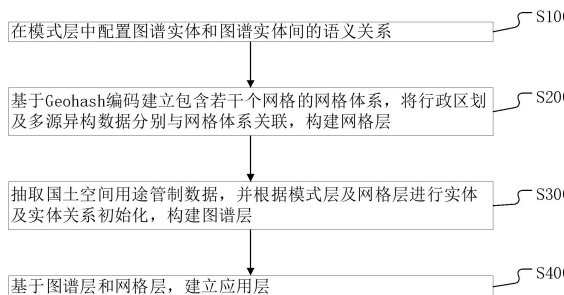
权利要求书4页 说明书19页 附图3页

(54) 发明名称

一种基于时空知识图谱的国土空间用途管制数据处理方法

(57) 摘要

本申请涉及一种基于时空知识图谱的国土空间用途管制数据处理方法。所述方法包括：在模式层中配置图谱实体和所述图谱实体间的语义关系，图谱实体的类别包括用于表征时空切片状态的时空实体，以及用于表征对土地状态产生实际或潜在影响的人类活动的事件实体；基于Geohash编码建立包含若干个网格的网格体系，将行政区划及多源异构数据分别与网格体系关联，构建网格层；抽取国土空间用途管制数据，并根据模式层及网格层进行实体及实体关系初始化，构建图谱层；基于所述图谱层和所述网格层，建立应用层。相较于现有技术，本申请可支持对地理现象的推理和预测，引入多维度信息特征，有效支撑业务应用。



1. 一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,包括:

在模式层中配置图谱实体和所述图谱实体间的语义关系;所述图谱实体的类别包括用于表征时空切片状态的时空实体,以及用于表征对土地状态产生实际或潜在影响的人类活动的事件实体;

基于Geohash编码建立包含若干个网格的网格体系,将行政区划及多源异构数据分别与所述网格体系关联,构建网格层;其中,所述多源异构数据包括栅格数据、矢量数据和POI数据中的至少一种;

抽取国土空间用途管制数据,并根据所述模式层及所述网格层进行实体及实体关系初始化,构建图谱层;

基于所述图谱层和所述网格层,建立应用层。

2. 根据权利要求1所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,所述时空实体 $Entity_{TS}$ 配置有时间、空间及属性,定义为:

$$Entity_{TS} = \{E_{id}, E_{time}, E_{space}, E_{QA}, E_{CA}\}$$

其中, E_{id} 表示实体编码, E_{time} 表示时空实体对应的时间, E_{space} 表示时空实体对应的空间, E_{QA} 表示定量属性, E_{CA} 表示类别属性;

所述事件实体 $Entity_E$ 配置有类型、时间、空间和影响范围,定义为:

$$Entity_E = \{E_{id}, E_{type}, E_{time}, E_{space}, E_{Scope}\}$$

其中, E_{id} 表示实体编码, E_{type} 表示事件实体的类型, E_{time} 表示事件发生的时间, E_{space} 表示事件发生的空间, E_{Scope} 表示事件影响的面积;

所述语义关系的类别包括用于表征网格与周围八邻域网格的双向关系的相邻关系、用于表征事件对象在历史的时空网格上的实施行为的发生关系、用于表征事件对象对未来时空网格的影响的影响关系、用于记录不同状态间时空网格的状态变化的过程关系,以及用于表征发生先后逻辑的顺承关系;其中,所述相邻关系为双向关系,所述发生关系、影响关系、过程关系及顺承关系为单向关系。

3. 根据权利要求2所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于:

所述发生关系 $Relationship_O$ 定义为:

$$Relationship_O = \{R_{id}, R_{type}, R_{time}\}$$

其中, R_{id} 表示关系编码, R_{type} 表示关系类型, R_{time} 表示事件发生时间;

所述发生关系与图谱实体间的关系定义为:

$$R_O = \{Entity_{TS}, Relationship_O, Entity_E\}$$

其中, $Entity_{TS}$ 为三元组的头节点,表示历史时空实体, $Entity_E$ 为三元组的尾节点;

所述影响关系 $Relationship_I$ 定义为:

$$Relationship_I = \{R_{id}, R_{type}, R_{time}\}$$

其中, R_{id} 表示关系编码, R_{type} 表示关系类型, R_{time} 表示事件影响时间;

所述影响关系与图谱实体间的关系定义为：

$$R_I = \{Entity_E, Relationship_I, Entity_{TS}\}$$

其中, $Entity_E$ 为三元组的头节点, 表示特定事件实体; $Entity_{TS}$ 为三元组的尾节点, 表示未来时空实体;

所述顺承关系 $Relationship_S$ 定义为：

$$Relationship_S = \{R_{id}, R_{type}\}$$

其中, R_{id} 表示关系编码, R_{type} 表示关系类型;

顺承关系与图谱实体间的关系定义为：

$$R_I = \{Entity_{E_A}, Relationship_S, Entity_{E_B}\}$$

其中, $Entity_{E_A}$ 为三元组的头节点, 表示前一阶段事件实体; $Entity_{E_B}$ 为三元组的尾节点, 表示后一阶段事件实体;

所述过程关系 $Relationship_P$ 定义为：

$$Relationship_P = \{R_{id}, R_{type}, R_{AC}\}$$

其中, R_{id} 表示关系编码, R_{type} 表示关系类型, R_{AC} 表示属性变化情况;

所述过程关系与图谱实体间的关系定义为：

$$R_I = \{Entity_{TS_A}, Relationship_P, Entity_{TS_B}\}$$

其中, $Entity_{TS_A}$ 为三元组的头节点, 表示前一阶段时空实体; $Entity_{TS_B}$ 为三元组的尾节点, 表示后一阶段事件实体;

所述相邻关系 $Relationship_A$ 定义为：

$$Relationship_A = \{R_{id}, R_{type}, R_{distance}, E_{direction}\}$$

其中, R_{id} 表示关系编码, R_{type} 表示关系类型, $R_{distance}$ 表示相邻距离, $E_{direction}$ 表示相邻方向;

所述相邻关系与图谱实体间的关系定义为：

$$R_I = \{Entity_{TS}, Relationship_A, Entity_{TS_S}\}$$

其中, $Entity_{TS}$ 为三元组的头节点, 表示特定时空实体, $Entity_{TS_S}$ 为三元组的尾节点, 表示相邻时空实体。

4. 根据权利要求2所述的一种基于时空知识图谱的国土空间用途管制数据处理方法, 其特征在于, 所述构建网格层, 包括:

利用Geohash编码将单位区域的二维经纬度转换为一维字符串编码, 形成配置有网格编码的Geohash网格单元;

将行政区域代码与相应区域的所述Geohash网格单元的所述网格编码关联, 建立行政区域边界与Geohash网格的拓扑关系, 生成网格-行政区关联表;

获取每个网格对应的四至信息, 并将所述四至信息重构为矢量格网;

基于所述栅格数据获取每个网格对应的波段信息, 并将所述波段信息赋予所述矢量格网后, 将所述矢量格网与相应网格进行叠加, 实现所述栅格数据的网格化;

将所述矢量数据与网格进行叠加,确定唯一图斑,并排除拓扑错误,实现矢量数据网格化;

将所述POI数据中的POI点与网格进行挂接,实现POI数据网格化;

根据所述网格-行政区关联表,基于所述Geohash网格单元所关联的行政区域的中心点坐标,为所述Geohash网格单元的各个图斑定义投影坐标;

基于网格化的栅格数据、矢量数据和POI数据,结合所定义的投影坐标进行空间叠置分析,并基于Geohash编码实现空间图层的融合。

5. 根据权利要求4所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,所述构建图谱层,包括:

基于Geohash编码和哈希码定义编码体系;其中,所述编码体系中的各个统一编码映射有概念类别、相应概念类别下的类型、时间及Geohash网格单元;

建立实体库表;所述实体库表包括时空实体表及事件实体表,分别用于存储所述时空实体及所述事件实体的参数信息;

建立关系库表,用于存储所述图谱实体间的相应语义关系,包括发生关系表、影响关系表、过程关系表、相邻关系表及顺承关系;其中,所述发生关系表、影响关系表和顺承关系表记录有相应的语义关系的类型,及头节点、尾节点的统一编码及时间节点;所述相邻关系表记录有相应的语义关系的类型,头节点、尾节点的统一编码和时间节点,以及对应网格间的距离;所述过程关系表记录有相应的语义关系的类型,头节点、尾节点的统一编码和时间节点,以及不同时间节点的业务属性变化情况;

对所述国土空间用途管制数据进行实体属性抽取,完成所述时空实体及所述事件实体的初始化,并基于初始化后的所述时空实体和事件实体进行实体关系初始化,以为所述实体库表和所述关系库表的初始化提供数据;其中,所述实体关系初始化包括空间关联、业务关联和时态关联;

初始化所述发生关系表、所述影响关系表、所述顺承关系表、所述过程关系表和所述相邻关系表;

结合所定义的所述编码体系,对初始化后的所述时空实体、所述事件实体、所述发生关系表、所述影响关系表、所述顺承关系表、所述过程关系表和所述相邻关系表进行编码,形成图谱层数据。

6. 根据权利要求5所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,所述空间关联包括:

以Geohash编码作为连接键,基于从所述国土空间用途管制数据抽取出的历史事件信息及业务逻辑,为初始化的时空实体挂接表征历史事件信息的事件实体,以为后续所述发生关系表、所述影响关系表及所述过程关系表的初始化提供数据;

基于所述国土空间用途管制数据,记录相邻的时空实体,以为后续所述相邻关系表的初始化提供数据;

所述业务关联包括:

根据所述业务逻辑将对应的事件实体关联至另一事件实体,以为后续所述顺承关系表的初始化提供数据;

所述时态关联包括:

根据预设的时间聚合单元,对所述事件实体进行聚合,并与对应的Geohash网格单元关联,以为后续所述顺承关系表的初始化提供数据。

7.根据权利要求6所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,所述业务关联还包括:

从所述国土空间用途管制数据中抽取出管理信息,并根据所述管理信息为对应的时空实体关联对应的管控指标。

8.根据权利要求4-7任一项所述的一种基于时空知识图谱的国土空间用途管制数据处理方法,其特征在于,所述网格层还配置有复合指标、景观指数指标及距离指标中的至少一类,以及用于实现相应指标计算的网格算法。

9.一种计算机可读存储介质,其特征在于,所述存储介质上存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行,以实现如权利要求1-8任一项所述方法。

10.一种计算机程序产品,包括计算机程序或计算机可执行指令,其特征在于,所述计算机程序或计算机可执行指令被处理器执行时,实现权利要求1-8任一项所述方法。

一种基于时空知识图谱的国土空间用途管制数据处理方法

技术领域

[0001] 本申请涉及地理信息数据处理技术领域,特别是涉及一种基于时空知识图谱的国土空间用途管制数据处理方法。

背景技术

[0002] 国土空间用途管制作作为自然资源管理的重要内容,是政府为了保护国土空间资源,确保领土、领海和领空被合理保护、开发和利用而采取的一系列法律、行政和技术措施,用以限制或指导国土的使用和转换。如耕地用途管制,其涉及耕地保护目标、永久基本农田、耕地后备资源管理、耕地质量分类等多类业务数据,且业务数据之间存在复杂的业务链条和业务逻辑,在耕地监管、耕地布局优化、遏制违法占耕等场景中,往往需要使用多类耕地用途管制数据协同分析才可获取有效知识,然而现有技术难以支撑耕地用途管制数据有效关联和协同应用。

[0003] 现有技术主要利用自然语言处理方法发现知识中存在的关联关系,进而构建知识图谱,实现数据关联和处理,但该模式在自然资源管理领域存在三大不足:

首先,国土空间用途管制所涉业务数据间的关联关系存在相应的业务逻辑,这些业务逻辑由业务专家制定,若直接采用自然语言处理方式构建知识图谱会导致知识建模不全面、不系统、更新困难等问题,难以支撑业务应用。

[0004] 其次,自然资源数据不仅包含语义信息,还涉及复杂的时空特征和地学机理,如地理实体的空间分布、时间序列上的演变规律、以及地理过程的动态变化等,通用知识图谱难以表达以上多维度信息特征。

[0005] 最后,通用知识图谱在处理动态变化数据时表现不足,无法完整呈现地理过程的时空演变规律,也难以支持对地理现象的推理和预测。

发明内容

[0006] 基于此,有必要针对上述的至少一种技术问题,提供一种基于时空知识图谱的国土空间用途管制数据处理方法。

[0007] 为解决上述技术问题,本发明的技术方案如下:

第一方面,一种基于时空知识图谱的国土空间用途管制数据处理方法,包括:

在模式层中配置图谱实体和所述图谱实体间的语义关系;所述图谱实体的类别包括用于表征时空切片状态的时空实体,以及用于表征对土地状态产生实际或潜在影响的人类活动的事件实体;

基于Geohash编码建立包含若干个网格的网格体系,将行政区划及多源异构数据分别与所述网格体系关联,构建网格层;其中,所述多源异构数据包括栅格数据、矢量数据和POI数据中的至少一种;

抽取国土空间用途管制数据,并根据所述模式层及所述网格层进行实体及实体关系初始化,构建图谱层;

基于所述图谱层和所述网格层,建立应用层。

[0008] 第二方面,一种计算机可读存储介质,所述存储介质上存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行,以实现第一方面所述方法。

[0009] 第三方面,一种计算机程序产品,包括计算机程序或计算机可执行指令,所述计算机程序或计算机可执行指令被处理器执行时,实现第一方面所述方法。

[0010] 与现有技术相比,本发明技术方案的有益效果是:

本申请公开了一种基于时空知识图谱的国土空间用途管制数据处理方法,针对国土空间用途管制数据关联难的问题,基于“状态-事件-过程”的时空知识图谱组织模式,综合考虑业务逻辑、地学机理和时空演变过程,设计了国土空间用途管制时空知识图谱概念模型(即模式层),进而支持对地理现象的推理和预测;利用Geohash编码构建空间基础网格,实现空间的网格化和数据的网格化,从而引入多维度信息特征;在数据网格化基础上,实现基于时空知识图谱的国土空间用途管制数据处理,可有效支撑业务应用层(如描述型应用、计算型应用、诊断型应用等)。

附图说明

[0011] 图1为本申请一些实施例中数据处理方法的流程示意图。

[0012] 图2为本申请一些实施例中时空知识图谱的模型示意图。

[0013] 图3为本申请一些实施例中基于时空知识图谱的查询结果示意图。

[0014] 图4为本申请一些实施例中实体关联结果示意图。

[0015] 图5为本申请一些实施例中数据处理方法的另一流程示意图。

具体实施方式

[0016] 本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换,这仅仅是描述本申请的实施例中对相同属性的对象在描述时所采用的区分方式。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,以便包含一系列单元的过程、方法、系统、产品或设备不必限于那些单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它单元。术语“确定”广泛涵盖各种各样的动作,可包括获取、演算、计算、处理、推导、调研、查找(例如,在表、数据库或其他数据结构中查找)、探明、和类似动作,还可包括接收(例如,接收信息)、访问(例如,访问存储器中的数据)和类似动作,还可包括生成、创建、建立和类似动作,以及解析、选择、选取和类似动作等等。其他术语的相关定义将在下文描述中给出。

[0017] 需要说明的是,当一个元件被认为是“连接”另一个元件时,它可以是直接连接到另一个元件,或者通过居中元件连接另一个元件。此外,以下实施例中的“连接”,如果被连接的对象之间具有电信号或数据的传递,则应理解为“电连接”、“通信连接”等。

[0018] 需要强调的是,本申请实施例技术方案中对数据的获取、传输、存储、使用、处理等均符合国家法律法规的相关规定。

[0019] 在本申请实施例中,可能提及某些软件、组件、模型等业界已有方案,应当将它们

认为是示范性的,其目的仅仅是为了说明本申请技术方案实施中的可行性,但并不意味着申请人已经或者必然用到了该方案。

[0020] 附图仅用于示例性说明,不能理解为对本专利的限制;

为了更好地说明本实施例,附图某些部件会有省略、放大或缩小,并不代表实际产品的尺寸;

对于本领域技术人员来说,附图中某些公知结构及其说明可能省略是可以理解的。

[0021] 下面结合附图和实施例对本发明的技术方案做进一步的说明。

[0022] 需要说明的是,时空知识图谱是通用知识图谱与地理学结合的产物,是具有地理时空分布或位置隐喻的知识构成的有向图,即以时空分布特征为核心的知识图谱。

[0023] 在本申请的一些实施例中引入时空知识图谱理论体系,提供一种基于时空知识图谱的国土空间用途管制数据处理方法,参阅图 1,包括步骤S100至步骤S400。

[0024] S100、在模式层中配置图谱实体和所述图谱实体间的语义关系(本申请中也简称“实体关系”);所述图谱实体的类别包括用于表征时空切片状态的时空实体,以及用于表征对土地状态产生实际或潜在影响的人类活动的事件实体;

S200、基于Geohash编码建立包含若干个网格的网格体系,将行政区划及多源异构数据分别与所述网格体系关联,构建网格层;其中,所述多源异构数据包括栅格数据、矢量数据和POI(Point of Interest,兴趣点)数据中的至少一种;

S300、抽取国土空间用途管制数据,并根据所述模式层及所述网格层进行实体及实体关系初始化,构建图谱层。

[0025] S400、基于所述图谱层和所述网格层,建立应用层。

[0026] 上述实施例中,首先,构建了国土空间用途管制时空知识图谱模式层,用于配置图谱的核心概念(本领域技术人员也称为“本体”)和语义关系。时空知识图谱模式层包括时空网格和驱动事件两个核心本体(即所述时空本体与事件实体)。其中,时空网格是时空知识图谱中承载实体信息的基础单元,具有时间、空间、属性等特征维度,通过这三个维度构建“时空实体”,其本体状态为该网格在某一时间点上土地利用、环境特征、累积历史业务的综合描述。驱动事件用于表征对土地状态产生实际或潜在影响的人类活动,主要包括土地用途管制类活动和土地开发类活动,土地用途管制活动包括划定永久基本农田、发布土地规划用途管制政策等;土地开发活动包括建设用地审批、垦造水田项目实施等,驱动事件具有事件类型、发生时间、影响面积等属性,这些事件的发生不仅影响当下的土地状态,也可对未来状态产生累积性效应,通过时间、空间、事件类型、影响范围四个维度构建“事件实体”。通过构建时空实体与事件实体之间的语义关系,可以清晰地表达土地利用状态的演变路径,即同一格网在不同时间节点的属性变化过程,为国土空间用途管制提供更精准的信息支持,有助于更好地理解 and 把握国土空间利用的动态变化规律。

[0027] 其次,网格层则引入Geohash空间编码剖分空间,实现空间数据标准化,用以支撑图谱实体及语义关系的实例化。作为非限制性示例,所述网格层采用PostgreSQL数据库存储。

[0028] 在一些具体实施过程中,所述网格体系为将空间均匀剖分的网格体系。

[0029] 最后,在数据网格化基础上,基于国土空间用途管制数据,结合时空知识图谱概念

模型(即模式层),通过实体及实体关系初始化,为剖分的空间赋予时间属性,实现时空标准化表征,构建面向国土空间用途管制的时空知识图谱概念模型,从而有效整合国土空间用途管制数据,实现数据有效关联处理和协同应用。相较于现有技术,上述实施例解决了国土空间利用状态难以表达、事件链条难以组织、风险行为难以识别等问题,从而支撑国土空间用途管制数据协同应用,增强国土空间保护监管智能化水平。

[0030] 可以理解,上述实施例中,将国土空间相关的时空数据(对应时空实体)、事件数据(对应事件示例)等多种数据进行了有效的整合,从而构建时空知识图谱,实现数据的统一存储和关联,便于数据的查询、分析和共享,提高了数据的利用效率和价值。

[0031] 在一些具体实施过程中,所述模式层采用PostgreSQL数据库进行存储。

[0032] 在一些优选实施例中,涉及步骤S100,所述时空实体 $Entity_{TS}$ 配置有时间、空间及属性,定义为:

$$Entity_{TS} = \{E_{id}, E_{time}, E_{space}, E_{QA}, E_{CA}\}$$

[0033] 其中, E_{id} 表示实体编码(此处为时空实体的编码), E_{time} 表示时空实体对应的时间, E_{space} 表示时空实体对应的空间, E_{QA} 表示定量属性(如耕地面积、林地面积、城镇开发边界面积等), E_{CA} 表示类别属性(如土壤质地等级、主体功能区类别等);

所述事件实体 $Entity_E$ 配置有类型、时间、空间和影响范围,定义为:

$$Entity_E = \{E_{id}, E_{type}, E_{time}, E_{space}, E_{Scope}\}$$

[0034] 其中, E_{id} 表示实体编码(此处为事件实体的编码), E_{type} 表示事件实体的类型, E_{time} 表示事件发生的时间, E_{space} 表示事件发生的空间, E_{Scope} 表示事件影响的面积;

所述语义关系的类别包括用于表征网格与周围八邻域网格的双向关系的相邻关系、用于表征事件对象在历史的时空网格上的实施行为的发生关系、用于表征事件对象对未来时空网格的影响的影响关系、用于记录不同状态间时空网格的状态变化的过程关系,以及用于表征发生先后逻辑的顺承关系;其中,所述相邻关系为双向关系,所述发生关系、影响关系、过程关系及顺承关系为单向关系。

[0035] 应当理解,上述优选实施例中,定义了“相邻关系”、“发生关系”、“影响关系”、“过程关系”、“顺承关系”(简称“相邻”、“发生”、“影响”、“过程”、“顺承”)五类时空知识图谱的语义关系;其中,“相邻关系”定义了网格单元与八邻域网格的双向关联,“发生关系”表征事件对象在历史时空网格上的实施行为,“影响关系”表征事件对象对未来时空网格的影响,“过程关系”记录不同状态间时空网格属性变化;“顺承关系”定义事件发生先后逻辑。

[0036] 如图 2所示,橙色矩形为相同位置上不同时间节点的时空实体,具备时空定位信息与属性信息;绿色多边形为土地开发或管制事件,可通过叠置分析将事件属性与网格关联,黑色箭头表示语义关系,包括“相邻”“发生”“影响”“过程”“顺承”五类关系,其中“相邻关系”为双向关系,用双箭头表达,其余为单向关系。通过描述地理现象“历史状态-事件-未来状态”的概念模式,完成国土空间用途管制时空知识图谱模型设计。

[0037] 作为非限制性示例,“相邻关系”为通过Geohash编码的空间邻近算法,支持表征空间邻域效应。

[0038] 在一些优选实施例中,所述发生关系 $Relationship_O$ 定义为:

$$Relationship_O = \{R_{id}, R_{type}, R_{time}\}$$

[0039] 其中, R_{id} 表示关系编码(此处为发生关系的编码), R_{type} 表示关系类型, R_{time} 表示事件发生时间;

所述发生关系与图谱实体间的关系定义为:

$$R_O = \{Entity_{TS}, Relationship_O, Entity_E\}$$

[0040] 其中, $Entity_{TS}$ 为三元组的头节点,表示历史时空实体, $Entity_E$ 为三元组的尾节点;

所述影响关系 $Relationship_I$ 定义为:

$$Relationship_I = \{R_{id}, R_{type}, R_{time}\}$$

[0041] 其中, R_{id} 表示关系编码(此处为影响关系的编码), R_{type} 表示关系类型, R_{time} 表示事件影响时间;

所述影响关系与图谱实体间的关系定义为:

$$R_I = \{Entity_E, Relationship_I, Entity_{TS}\}$$

[0042] 其中, $Entity_E$ 为三元组的头节点,表示特定事件实体; $Entity_{TS}$ 为三元组的尾节点,表示未来时空实体;

所述顺承关系 $Relationship_S$ 定义为:

$$Relationship_S = \{R_{id}, R_{type}\}$$

[0043] 其中, R_{id} 表示关系编码(此处为顺承关系的编码), R_{type} 表示关系类型; 顺承关系与图谱实体间的关系定义为:

$$R_I = \{Entity_{E_A}, Relationship_S, Entity_{E_B}\}$$

[0044] 其中, $Entity_{E_A}$ 为三元组的头节点,表示前一阶段事件实体; $Entity_{E_B}$ 为三元组的尾节点,表示后一阶段事件实体;

所述过程关系 $Relationship_P$ 定义为:

$$Relationship_P = \{R_{id}, R_{type}, R_{AC}\}$$

[0045] 其中, R_{id} 表示关系编码(此处为过程关系的编码), R_{type} 表示关系类型, R_{AC} 表示属性变化情况;

所述过程关系与图谱实体间的关系定义为:

$$R_I = \{Entity_{TS_A}, Relationship_P, Entity_{TS_B}\}$$

[0046] 其中, $Entity_{TS_A}$ 为三元组的头节点,表示前一阶段时空实体; $Entity_{TS_B}$ 为三元组的尾节点,表示后一阶段事件实体;

所述相邻关系 $Relationship_A$ 定义为:

$$Relationship_A = \{R_{id}, R_{type}, R_{distance}, E_{direction}\}$$

[0047] 其中, R_{id} 表示关系编码(此处为相邻关系的编码), R_{type} 表示关系类型, $R_{distance}$ 表示相邻距离, $E_{direction}$ 表示相邻方向;

所述相邻关系与图谱实体间的关系定义为:

$$R_I = \{Entity_{TS}, Relationship_A, Entity_{TS_S}\}$$

[0048] 其中, $Entity_{TS}$ 为三元组的头节点,表示特定时空实体, $Entity_{TS_S}$ 为三元组的尾节

点,表示相邻时空实体。

[0049] 在一些优选实施例中,所述构建网格层,包括:

S2010、利用Geohash编码将单位区域的二维经纬度转换为一维字符串编码,形成配置有网格编码的Geohash网格单元。

[0050] 在一些具体实施过程中,采用7位Geohash编码将目标区域二维经纬度转换为一维字符串编码,对应空间精度约为 $152\text{m} \times 152\text{m}$,编码规则遵循RFC 4639 标准,确保空间邻近性与编码前缀的一致性。

[0051] 作为非限制性示例,将北纬 28.2° 、东经 112.9° 转换为 Geohash 编码“wpm4h7q”,对应约 $152.9\text{m} \times 152.4\text{m}$ 的网格单元。

[0052] S2020、将行政区域代码与相应区域的所述Geohash网格单元的所述网格编码关联,建立行政区域边界与Geohash网格的拓扑关系,生成网格-行政区关联表。

[0053] 在一些具体实施过程中,利用geopandas的sjoin空间连接方法,建立村级行政区边界与Geohash网格的拓扑关系,将行政区代码与覆盖该区域的网格编码关联,生成网格-行政区关联表。

[0054] S2030、获取每个网格对应的四至信息,并将所述四至信息重构为矢量格网。

[0055] 作为非限制性示例,利用gdal模块的GetGeoTransform()接口获取每个栅格的四至信息,并利用shapely.geometry库将四至信息重构为矢量格网。

[0056] S2040、基于所述栅格数据获取每个网格对应的波段信息,并将所述波段信息赋予所述矢量格网后,将所述矢量格网与相应网格进行叠加,实现所述栅格数据的网格化。

[0057] 在一些具体实施过程中,所述栅格数据为遥感影像数据。

[0058] 作为非限制性示例,利用gdal模块的ReadAsArray()接口获取每个栅格的波段信息,将波段信息赋予矢量格网。接着使用geopandas中overlay()接口实现栅格数据与网格的叠置分析。

[0059] S2050、将所述矢量数据与网格进行叠加,确定唯一图斑,并排除拓扑错误,实现矢量数据网格化。

[0060] 作为非限制性示例,以EPSG:4490为基准坐标系,采用geopandas中的overlay()接口进行外部数据和空间网格的叠加。

[0061] 在一些具体实施过程中,若某属性面积超过网格理论面积(如设定为 $152.9\text{m} \times 152.4\text{m} \approx 23302\text{m}^2$),则说明存在重叠图斑,检查重叠图斑业务属性是否一致,若一致,按面积字段去重,保留唯一图斑;若不一致,但仍由于图斑重叠导致面积超出格网面积,将面积修正为网格理论面积。

[0062] 在一些具体实施过程中,对于存在属性缺失的图斑,若为连续型列缺失值,则统一填充0值;若为分类列缺失值,则采用本列众数填充。

[0063] 在另一些具体实施过程中,先通过geopandas的explain_validity()方法检查图斑是否存在拓扑错误,再使用geopandas中的buffer(0)函数接口,修正图斑自相交、边界线重叠、内部孔洞与外边界存在共边等问题,为数据叠置分析做准备。

[0064] S2060、将所述POI数据中的POI点与网格进行挂接,实现POI数据网格化。

[0065] 作为非限制性示例,利用geopandas中的sjoin()接口将POI点与网格矢量挂接,并汇总统计网格内不同类型POI点的数量和属性值。

[0066] S2070、根据所述网格-行政区关联表,基于所述Geohash网格单元所关联的行政区域的中心点坐标,为所述Geohash网格单元的各个图斑定义投影坐标。

[0067] 在一些具体的实施过程中,基于网格关联的县级行政区中心点坐标,为其定义对应的投影坐标系,定义规则如下:

若中心点坐标经度范围在 $109.5^{\circ}\text{E} \sim 112.5^{\circ}\text{E}$ 之间,则定义投影坐标系为EPSG:4525;

若中心点坐标经度范围在 $112.5^{\circ}\text{E} \sim 115.5^{\circ}\text{E}$ 之间,则定义投影坐标系为EPSG:4526;

若中心点坐标经度范围在 $115.5^{\circ}\text{E} \sim 118.5^{\circ}\text{E}$ 之间,则定义投影坐标系为EPSG:4527。

[0068] 应当理解,通过为每个图斑定义投影,从而支持空间叠置分析,实现图斑面积、周长的精确计算。

[0069] S2080、基于网格化的栅格数据、矢量数据和POI数据,结合所定义的投影坐标进行空间叠置分析,并基于Geohash编码实现空间图层的融合。

[0070] 在一些具体实施过程中,在定义坐标系基础上,利用geopandas的area和length属性,获取每一个叠置图斑的面积,利用该方法实现面积、周长等属性的准确计算。

[0071] 需要说明的是,由于空间叠置分析后,指标数据图层处于“打散”状态,即数据分散在不同的图斑上,每个图斑只包含部分指标信息。因此需要将同一Geohash格网的图斑数据汇总,计算每个网格统一的指标值。

[0072] 在一些具体实施过程中,基于空间唯一的Geohash编码,利用Geopandas的dissolve()函数,实现空间图层的融合,利用dissolve()函数的aggfunc属性指定属性汇总方法,如加和计算指定汇总属性为“sum”,均值运算指定汇总属性为“mean”,从而实现网格融合。

[0073] 此外,在网格融合过程中,由于数据类型不同,需要采用不同汇聚方式实现属性聚合。

[0074] 在一些具体实施过程中,若数据类型为常量(如人口数量、建设用地总面积),在网格融合时,直接进行图斑属性算数加和处理;

若数据类型为当量(如土壤污染等级,土壤有机质含量等级),在网格融合时,按照图斑面积占网格面积比加权求和方式处理;

若数据类型为类别型(主体功能区类型),在网格融合时,取面积最大的图斑指标。

[0075] 在一些优选实施例中,所述网格层还配置有复合指标、景观指数指标及距离指标中的至少一类。

[0076] 在一些具体实施过程中,通过网格算法实现复合指标、景观指数指标及距离指标的配置,从而提供更为丰富的时空知识服务。

[0077] 作为非限制性示例,用于复合指标计算的网格算法有:

(1) 网格变化量算法

筛选具有相同Geohash编码,且时间相邻的网格,利用pandas库的merge()方法合并两个网格需要计算的属性,计算变化量和变化率,将计算结果返回原始网格数据。

[0078] (2) 指标加权算法

根据业务分析,选择参与评价的指标和指标权重,根据设计的正负向指标采用不同的方式进行标准化,然后利用熵权法计算每个指标的权重,最后利用numpy中的dot()方法计算复合指标值。

[0079] (3) 指标分级算法

首先利用IQR方法排除指标异常值,然后利用sklearn.preprocessing库中的StandardScaler类实现数据的标准化,利用sklearn.cluster库中的KMeans聚类方法实现数据的分级,利用肘部法则确定最佳聚类数,根据指标导向,按照每一列的指标均值对分级结果进行排序,最终得到指标分级结果。

[0080] 作为非限制性示例,所述景观指数指标包括网格景观指数和网格密度指标;其中,用于网格景观指数计算的网格算法有:

(1) 景观类型丰富度计算:景观类型丰富度即每个网格内包含的土地利用类型数量。首先构造土地利用类型表,然后遍历网格,利用pandas库的nunique()方法计算每个网格内土地利用类型数量,从而实现网格景观指数计算。

[0081] (2) 景观破碎度计算:如,通过计算网格内斑块数量与网格面积的比值(反映景观被分割的程度,值越高越破碎)确定。遍历网格,统计网格内斑块数量,利用网格内斑块数量除以网格面积。

[0082] (3) 香农多样性计算:反映景观类型的丰富度与分布均匀性,值越高表示多样性越高。遍历网格,首先计算网格内每一地类的数量与地类总数的占比,然后再计算该占比的信息熵,进而得到香农多样性指数。

[0083] 香农多样性指标计算公式如下:

$$H' = - \sum_{i=1}^n p_i * \ln(p_i)$$

[0084] 其中, H' 是香农多样性指数, p_i 是第*i*个景观类型的相对丰度。

[0085] 用于网格密度指标计算的网格算法为:遍历网格,计算对应指标与网格总面积之比。

[0086] 作为非限制性示例,距离指标包括网格距离、设施服务范围、网格服务效能等。

[0087] 其中,用于网格距离指标计算的网格算法(也称“网格距离算法”)可以通过GeoHash 编码的空间邻接特性,以圈层扩展方式计算网格到目标点的最少外扩圈层数来实现,包括:

(1) 邻域关系构建方法。利用pygeohash.get_adjacent()方法,为每个网格的GeoHash 生成 8 个相邻区域(包括上下左右及四个对角线方向),构建邻域关系表。

[0088] (2) 圈层距离计算方法。对每个目标网格,从自身Geohash出发(第0层),逐层向外扩展邻域,第一层包含目标网格的8个直接邻域;第二层包含第1层所有网格的8个邻域;以此类推,最多扩展20层。若某个圈层中存在目标点的Geohash,则该圈层序号为网格到目标点的距离;若扩展 20 层后未找到目标点,则距离记为-1。

[0089] 以及,用于设施服务范围的网格算法(也称“设施服务范围算法”)可以基于设施的GeoHash和服务距离(圈层数),通过圈层扩展标记所有可达网格来实现,包括:

(1) 服务范围初始化。为每个网格生成in_service_area字段,初始值为 0(不在服

务范围内)。

[0090] (2) 圈层扩展标记。对每个设施,从其 GeoHash 出发(第0层),按服务距离扩展圈层:第0层:设施所在网格,标记为1;第1层:设施的8个邻域网格,标记为1;依此类推,直到达到服务距离指定的圈层数。所有被标记为1的网格即属于该设施的服务范围。

[0091] 以及,用于网格服务效能计算的网格算法(也称“网格服务效能算法”)可以结合设施服务范围与目标对象属性,统计服务范围内符合条件的对象数量来实现,包括:

(1) 服务能力初始化。为每个设施生成service_capability字段,初始值为 0。

[0092] (2) 服务范围属性筛选。对每个设施,按服务距离扩展圈层,生成服务范围的 GeoHash 集合;在目标对象数据中,筛选出满足以下条件的对象:GeoHash属于服务范围集合;属性类型与设施服务对象匹配。符合条件的对象数量即为该设施的服务能力。

[0093] 可以理解,基于该网格计算思路,将传统的几何距离转化为 GeoHash 圈层差,避免复杂空间计算;利用get_adjacent()直接获取邻域,时间复杂度从 $O(n^2)$ 降至 $O(1)$;通过调整圈层数参数,可灵活处理具有不同服务范围的设施服务能力计算。

[0094] 在一些优选实施例中,所述构建图谱层,包括步骤S3010至S3060。

[0095] S3010、基于Geohash编码和哈希码定义编码体系;其中,所述编码体系中的各个统一编码映射有概念类别、相应概念类别下的类型、时间及Geohash网格单元。

[0096] 在一些具体实施过程中,定义统一编码的编码结构为28位编码,[概念类(1位)]+[类型(2位)]+[时间(8位)]+[位置(8位)]+[哈希校验(9位)];

各部分的编码表示如下:

(1) 概念类别

图谱实体(S)、语义关系(G)

(2) 类型

实体类型:时空实体(00)、事件实体(01)

关系类型:相邻关系(00)、顺承关系(01)、发生关系(02)、影响关系(03)、过程关系

(04)

(3) 时间:YYYYMMDD 格式(如 20250608)

(4) 位置(即Geohash网格单元):Geohash 编码前8位(截取 7 位 Geohash 后补零至8位)

(5) 哈希校验:基于MD5生成9位哈希值,避免节点重复。

[0097] 基于该编码体系,实现了时空知识图谱要素类型、时间、空间的统一表达。

[0098] S3020、建立实体库表(简称“实体表”);所述实体库表包括时空实体表及事件实体表,分别用于存储所述时空实体及所述事件实体的参数信息。

[0099] 在一些具体实施过程中,所述时空实体表存储时间、空间、业务三个维度信息。其中,空间维度采用Geohash 编码作为空间唯一标识;时间维度以每年土地变更调查启用的12月31日作为时间节点,记录本年的业务信息,实现空间状态和业务活动历史演变追踪;业务维度则集成耕地面积、建设用地面积、人口数量等国土空间业务属性,实现空间状态与业务活动的统一存储。

[0100] 在一些具体实施过程中,所述事件实体表记录业务的位置、发生时间、业务类型、影响面积等属性。

[0101] S3030、建立关系库表(简称“关系表”),用于存储所述图谱实体间的相应语义关系,包括发生关系表、影响关系表、过程关系表、相邻关系表及顺承关系;其中,所述发生关系表、影响关系表和顺承关系表记录有相应的语义关系的类型,及头节点、尾节点的统一编码和时间节点;所述相邻关系表时记录有相应的语义关系的类型,头节点、尾节点的统一编码和时间节点,以及对应网格间的距离;所述过程关系表记录有相应的语义关系的类型,头节点(也称“首节点”)、尾节点的统一编码和时间节点,以及不同时间节点的业务属性变化情况。

[0102] 在一些具体实施过程中,以年为单位聚合时,所述相邻关系表中存储尾节点的时间节点与首节点的时间节点(即首节点对应实体被记载的时间)一致。

[0103] 需要说明的是,步骤S3020和S3030通过构建实体库表和关系库表,能够有效描述国土空间变化过程,为后续国土空间用途管制时空知识图谱搭建奠定基础。

[0104] S3040、对所述国土空间用途管制数据进行实体属性抽取,完成所述时空实体及所述事件实体的初始化,并基于初始化后的所述时空实体和事件实体进行实体关系初始化,以为所述实体库表和所述关系库表的初始化提供数据;其中,所述实体关系初始化(也称“实体关联”)包括空间关联、业务关联和时态关联。

[0105] 作为非限制性示例,在实体初始化过程中,将PostgreSQL数据库存储的属性、关系转换成Neo4j图数据库的存储形式。

[0106] 在一些具体实施过程中,实体初始化过程包括数据集成、数据质检、实体更新及生成用于指导初始化实体库表(即实体入库)的实体信息表;实体关系初始化过程包括空间关联、业务关联、时态关联及生成用于指导初始化关系库表的关系信息表。

[0107] 作为非限制性示例,所述国土空间用途管制数据可以是多年份自然资源调查监测成果、国土空间相关业务规则等。根据所构建的时空图谱本体,从多年份自然资源调查监测成果中抽取图斑面积、土地类型、土地用途等核心信息,开展数据集成与融合,通过实体属性抽取,形成包含空间坐落、权属、类型和用途属性等基本属性和时态属性的国土空间实体成果。在初始化过程中,若数据存在矛盾、冲突和不完善的情况,则进行数据更新,以确保国土空间实体信息的准确性和完整性,完成数据质检。

[0108] 以某地耕地实体为例,抽取出的实体属性如表 1、表 2所示。

[0109] 表 1耕地时空实体部分属性

| 属性名称 | 属性值 |
|-----------|----------|
| Geohash编码 | ws15c34 |
| 耕地面积 | 13121.25 |
| 建设用地面积 | 0.00 |
| 人口数量 | 573 |
| 坡度 | 0.57 |
| 时间 | 2019 |

[0110] 表 2耕地事件实体部分属性

| 属性名称 | 属性值 |
|------------|----------------|
| Geohash编码 | ws15c34 |
| 规划设施农用地面积 | 360.02 |
| 规划城镇建设用地面积 | 3965.75 |
| 事件名称 | 规划用地用海批复 |
| 事件时间 | 2022-09-05 |
| 事件id | Ws15c34ghdyhpf |

[0111] S3050、初始化所述发生关系表、所述影响关系表、所述顺承关系表、所述过程关系表和所述相邻关系表。

[0112] S3060、结合所定义的所述编码体系,对初始化后的所述时空实体、所述事件实体、所述发生关系表、所述影响关系表、所述顺承关系表、所述过程关系表和所述相邻关系表进行编码,形成图谱层数据。

[0113] 在一些具体实施过程中,对构造的时空知识图谱实体、关系初始化结果进行编码,形成CSV格式的实体表,利用Neo4j的LOAD CSV实现标准化入库。以Geohash编码作为关联键,构建发生关系、影响关系、相邻关系、顺承关系、过程关系。

[0114] 作为非限制性示例,以geohash为ws638hr的空间进行时间序列上的时空演化模式查询,查询结果如图 3所示。

[0115] 作为非限制性示例,所记录的事件时间可以精确到年、月、日,由本领域技术人员根据实际情况自行设定,本公开不对此做限制。

[0116] 在一些优选实施例中,所述方法还包括:

为图谱实体的关键属性添加约束和索引,提高数据查询效率。

[0117] 作为非限制性示例,为时空实体、事件实体的编码添加唯一约束,为实体的geohash编码、时间标签添加索引,以提升数据查询检索效率,实现时空知识图谱优化。

[0118] 在一些优选实施例中,所述空间关联包括:

以Geohash编码作为连接键,基于从所述国土空间用途管制数据抽取出的历史事件信息及业务逻辑,为初始化的时空实体挂接表征历史事件信息的事件实体,以为后续所述发生关系表、所述影响关系表及所述过程关系表的初始化提供数据;

基于所述国土空间用途管制数据,记录相邻的时空实体,以为后续所述相邻关系表的初始化提供数据。

[0119] 在一些具体实施过程中,基于实体初始化成果,以Geohash编码作为连接键,将时空单元(如国土格网对应时空实体)上发生的历史事件信息(如用地报批、土地供应等事件)挂接到时空实体。通过这一历史事件 - 时空实体挂接机制,补全历史用地报批信息、土地供应信息、土地使用信息、补充耕地信息等业务信息,建立历史事件与时空实体在空间维度的直接关联,描述历史事件对时空实体的累积效应,使时空实体具备更丰富的业务上下文。

[0120] 在一些具体实施过程中,对于相邻关系,除记录首结点、尾结点的基本标识信息外,还纳入格网相距距离等参数,精准刻画时空实体在空间上的邻接、分布特征,完成邻域关系量化,为后续空间分析(如区域特定类型土地变化联动分析)奠定基础。

[0121] 所述业务关联包括:

根据所述业务逻辑将对应的事件实体关联至另一事件实体,以为后续所述顺承关系表的初始化提供数据。

[0122] 在一些具体实施过程中,围绕国土空间保护核心业务逻辑,建立业务事件(即事件实体)与时空实体关联的规则与约束。

[0123] 作为非限制性示例,针对耕地资源,依据“占补平衡”制度,关联耕地占用事件与补充耕地事件,确保对耕地数量、质量平衡关系的有效跟踪与分析,支撑业务合规性校验与动态监管。

[0124] 所述时态关联包括:

根据预设的时间聚合单元,对所述事件实体进行聚合,并与对应的Geohash网格单元关联,以为后续所述顺承关系表的初始化提供数据。

[0125] 在一些具体实施过程中,以年为时间聚合单元,将不同业务发生时间按年进行聚合并与 Geohash 网格单元(对应时空实体)关联。通过该方式,记录 Geohash 网格单元全生命周期的属性变化(如耕地面积、森林覆盖率、土地类型变化)、业务变化(如经历的用地审批、规划调整等事件)的历史过程,为国土空间保护的历史信息回溯提供详实、可查的数据源。

[0126] 需要说明的是,由于以预设的时间聚合单元(如“年”)为时间记录单位,而业务事件(对应事件实体)是在一年内各个时间上发生的,也可能重复发生多次,所以Geohash网格单元(对应时空实体)在关联一年内发生的业务事件时,需要对该网格单元上发生的事件进行聚合。

[0127] 上述优选实施例,通过构造实体关系,建立与时空实体、事件实体映射,可描述具体空间位置的业务关联及时间变化。如图 4所示,按年创建时空实体以记载当年格网上的属性信息,按事件发生类型、事件发生时间创建空间规划、用地审批、土地供应等业务发生的事件实体,记载事件的重要属性。

[0128] 在一些优选实施例中,所述业务关联还包括:

从所述国土空间用途管制数据中抽取出管理信息,并根据所述管理信息为对应的时空实体关联对应的管控指标。

[0129] 作为非限制性示例,所述管理信息可以是国土空间实体对象在自然资源业务流过程中产生的各类信息,包括但不限于规划信息、占用土地开发信息、补充耕地信息、执法监督信息等。

[0130] 其中,规划信息可涵盖永久基本农田、土地保有量、建设用地规模等空间管控指标信息;占用土地开发信息有农转用面积、建设用地审批面积、单独选址项目面积等;补充耕地信息有耕地后备资源开发面积、补充耕地新增面积等;执法监督信息包含违法占用林地面积、耕地“非农化”面积等。

[0131] 通过将上述管理信息与对应时空实体进行关联,形成业务 - 指标 - 数据之间的结构化描述,为智能算法(如土地利用变化预测、违法占耕识别算法)挖掘数据潜在信息提供清晰、规范的输入,助力实现国土空间保护业务的智能化监管与决策。

[0132] 在一些优选实施例中,所述图谱层还配置有图谱算法。

[0133] 在一些具体实施过程中,所述图谱算法包括规则推理算法,具体如下:

(1) 业务规则抽象

对具体业务场景中异常判定条件进行形式化抽象,提炼出“实体属性条件+关系关联条件+异常事件结论”的规则模板,如表 3所示,具体包括属性异常和关系异常查询。

[0134] 表 3基于Cypher的查询模版示例

```
// 通用异常查询模板
MATCH {实体匹配子句}
WHERE {属性过滤子句}
OPTIONAL MATCH {关系遍历子句}
WHERE {关系属性过滤子句}
WITH {结果处理子句}
WHERE {最终过滤子句}
CREATE {结论生成子句}
RETURN {结果返回子句}
```

[0135] (2) 图谱元素映射

将抽象规则中的实体类型、属性条件、关系类型映射为知识图谱中的节点类型、属性字段及边关系。

[0136] (3) 查询语句构建

基于图数据库查询语言(如 Cypher),构建包含实体匹配、关系遍历、属性过滤的复合查询语句,实现规则逻辑的图谱化表达。

[0137] (4) 推理结果解析

执行查询语句后,对返回的图谱子图进行解析,提取满足规则条件的异常关联路径,生成推理结论。以耕地为例,属性异常查询和关系异常查询示例如表 4所示:

表 4异常查询示例

```

//属性异常查询示例
// 步骤1: 匹配耕地面积减少的时空实体
MATCH (s:时空实体)
WHERE s.t节点耕地面积 < s.(t-n)节点耕地面积 AND s.土地利用类型 = "耕地"
// 步骤2: 检查是否存在影响关系
OPTIONAL MATCH (s)-[:影响关系]-(e:事件实体)
WHERE e.类型 IN ["耕地征收", "建设占用", "灾害损毁"]
// 步骤3: 筛选无影响事件的异常情况
WITH s, count(e) as event_count
WHERE event_count = 0
// 步骤4: 推断并创建非农化事件
CREATE (newEvent:事件实体 {
  类型: "耕地非农化",
  推断时间: datetime(),
  影响面积: s.(t-n)节点耕地面积 - s.t节点耕地面积
})
CREATE (newEvent)-[:影响关系]->(s)
// 步骤5: 返回推理结果
RETURN s.geohash, s.耕地面积, newEvent

```

```

// 关系异常查询语句
MATCH (supply:事件实体 {类型: "供地"})
WHERE supply.面积 > 0 AND supply.面积 > (SELECT approve.面积 FROM :事件实体 approve WHERE
approve.类型 = "批地" AND approve.时间 < supply.时间)
OPTIONAL MATCH (supply)-[:顺承关系]->(approve:事件实体 {类型: "批地"})
WHERE approve.面积 >= supply.面积 AND approve.时间 < supply.时间
WITH supply, count(approve) as approve_count
WHERE approve_count = 0
CREATE (newEvent:事件实体 {
  类型: "未批先供",
  违规面积: supply.面积,
  供地时间: supply.时间
})
CREATE (newEvent)-[:顺承关系]->(supply)
RETURN supply.geohash, newEvent

```

[0138] 本领域技术人员应当理解,上述示例中,“(t-n)节点耕地面积”表示在t节点之前n个节点时的历史耕地面积。

[0139] 在一些具体实施过程中,所述图谱算法还包括时空推理算法,具体如下:

(1) 构建深度学习模型

构建由图神经网络(GNN)和Transformer两个组件构成的深度学习模块。其中GNN

用于建模实体间的空间关联与关系传递,在预定义的实体属性中,时空实体、事件实体等节点属性(如土地面积、事件类型、时间编码)可直接作为节点特征输入;边的关系类型、时间差、距离等可作为边特征。针对图谱中节点和边特征,GNN直接利用其中预定义的五类语义关系(相邻、发生、影响、顺承等)进行推理,例如通过相邻关系边,GNN聚合邻域节点的国土状态特征(如面积变化率、类型分布),计算空间扩散效应的影响权重,从而捕捉国土空间状态演变的复杂依赖。Transformer模块用于处理事件序列的长距离依赖关系,利用注意力机制,捕捉不同事件对未来事件发生的影响权重。例如,“批地事件”与“供地事件”之间的时序关联,以及早期发生的事件对当前耕地减少事件的影响程度,通过 Transformer 模块可自动学习并赋予相应的权重。

[0140] (2) 特征向量生成

将待预测的图谱数据输入GNN;图谱数据中的历史事件序列(如供地、审批用地等离散事件的类型与时间戳)编码为Transformer输入。待模块分别输出特征后,融合为统一的时空联合特征向量,输入全连接层进行二分类预测。

[0141] (3) 训练深度学习模型

将GNN、Transformer模块通过特征拼接层整合,构建端到端的时空推理预测模型。以历史数据为训练集,通过采用交叉熵损失函数训练模型参数,评估模型对二分类问题(如,是否发生耕地减少事件)的预测能力。最终使模型输出的预测事件(如耕地减少事件)概率与实际发生情况的误差最小化。

[0142] (4) 预测概率计算

模型通过前向传播计算,输出目标时空网格在未来特定时间段(如1年、3年)内发生特定类型土地(如,耕地、林地、草地、建筑用地等)减少事件的概率值,概率范围为[0, 1]。数值越高,表示事件发生的可能性越大。

[0143] (5) 结果分析与预警生成

若概率值超过设定阈值(如0.6),则判定为高风险,自动生成预警信息(如耕地面积过度减少);若在阈值范围内,则判定为中低风险。

[0144] 在一些具体实施过程中,所述图谱算法还包括因果推理算法,具体如下:

(1) 特征嵌入

图神经网络算法的特征嵌入具体通过调用PyTorch Geometric框架中的SAGEConv图卷积算子。该方法可用于处理地理空间特性,直接利用边定义的空间拓扑,无需强制规则化网格,并且在聚合邻域节点特征时注入空间关系权重作为动态聚合系数,生成包含空间关联信息的嵌入向量。

[0145] (2) CatBoost 模型架构设计

为利用上述嵌入向量的多层次信息(处理时空知识图谱中的异构特征),采用CatBoost梯度提升树架构作为下游预测模块。通过使用Catboost库中CatBoostClassifier类实现CatBoost(分类与回归树提升)算法来迭代构建回归树,每次迭代基于之前树的残差优化新树的生成,降低预测偏差。内置针对分类特征的最优分割策略,自动学习分类特征的有序性,避免手动编码带来的信息损失。

[0146] (3) 模型训练与优化

CatBoost初始时参数设置如下:

表 5 初始化参数设置示例

```

model = CatBoostClassifier(
    iterations=2000,
    learning_rate=0.1,
    depth=6,
    cat_features=cat_col, # 指定类别特征
    verbose=100,
    task_type="GPU", # 指定使用 GPU
    devices="0"
)

```

[0147] 采用网格搜索或随机搜索方法,优化CatBoost的关键参数,包括树的数量、学习率、叶节点数等,以验证集上的F1分数或AUC值为评价指标。

[0148] (4) 基于Shapley值的归因分析算法

通过调用shap库的TreeExplainer类进行shapley分析,其步骤包括:

1) 特征贡献度计算:对每个样本的每个特征,计算其Shapley值,反映该特征对样本预测结果(如违法占耕概率)的边际贡献。

[0149] 2) 全局重要性排序:对所有样本的特征Shapley值取均值,得到特征的全局重要性排序,识别对事件发生影响最大的核心因素。

[0150] 3) 局部解释生成:针对单个样本,通过Shapley值解释该样本预测结果的主要驱动特征,例如某违法占耕事件主要由“距城镇中心距离近”和“邻域工业用地占比高”导致。

[0151] 在一些具体实施过程中,所述图谱算法还包括图推理算法,具体如下:

(1) 社区发现算法

利用Louvain算法进行社区发现,以最大化模块度为目标,将时空图划分为若干社区。具体实现时,通过图数据库的Cypher查询语言构建时空图模型,利用图数据库对图结构数据的高效处理能力,快速执行Louvain算法迭代计算,实现社区划分。

[0152] (2) 构建土地利用剧烈变化区域筛选策略

设计土地变化剧烈度的多指标综合评分:

$$Score = \alpha \times EventDensity + \beta \times \Delta LU + \gamma \times Heterogeneity$$

[0153] 其中, $EventDensity$ 为事件密度,单位面积事件发生频次(次 / 平方公里),通过查询节点关联的发生关系边数量并结合时空网格面积计算; ΔLU 为利用变化强度,利用 $\sum_{t1}^{t2} \left| \frac{S_{t1} - S_{t2}}{S_{t1}} \right| \times 100\%$ 计算得到; $Heterogeneity$ 表示空间异质性,通过比较相邻节点的土地利用类型计算得到, α 、 β 、 γ 为权重系数,通过熵权法动态计算。

[0154] (3) 阈值确定方法

统计全量社区评分的均值 μ 与标准差 σ ,划定剧烈变化区域为评分大于 $\mu + 1.5\sigma$ 的社区。

[0155] 上述优选实施例中,从业务规则抽象到推理结果解析的规则推理过程,以及结合深度学习模型的时空推理等图谱算法,能够用于对国土空间用途管制中的异常情况进行检

测、预警和归因分析,如耕地非农化事件的推断、耕地减少事件的预测等,有助于加强国土空间用途管制,保障国土资源的合理利用和可持续发展。

[0156] 在一些优选实施例中,参阅图 5,涉及步骤S400,所述应用层包括描述型应用、计算型应用、诊断型应用及预测型应用中的至少一种。如土地利用演化过程的时空回溯、衍生知识的计算生成、业务链路异常检测、土地利用变化异常区域检测、异常现象发生因果推理、事件预测、属性预测和空间分布预测等,能够满足不同用户在国土空间用途管制及相关领域的多样化需求,提升时空知识服务的质量和水平。

[0157] 作为非限制性示例,所述描述型应用实现如下:

基于 Neo4j 图数据库的 Cypher 查询语言,按空间位置、时间区间、事件类型等多条件组合查询,实现土地利用演化过程的时空回溯。

[0158] 其中,空间范围查询如下:

表 6 空间范围查询示例

```
MATCH (s:时空实体) WHERE s.geohash STARTS WITH "GH34" RETURN s
```

[0159] 时间区间查询如下:

表 7 时间区间查询示例

```
MATCH (e:事件实体) WHERE e.发生时间 BETWEEN "2020-01-01" AND "2022-12-31" RETURN e
```

[0160] 关系链查询如下:

表 8 关系链查询示例

```
MATCH (s1:时空实体)-[:过程关系]->(s2:时空实体) RETURN s1, s2
```

[0161] 作为非限制性示例,所述计算型应用实现如下:

构建支持自定义指标配置的计算平台,基于网格层数据与空间分析算法,通过“基础指标提取 - 权重分配 - 模型计算”的流程生成衍生知识,并配置复合指标计算接口、景观指数指标计算接口、距离指标计算接口。

[0162] 其中,复合指标计算接口用于调用网格变化量算法、指标加权算法、指标分级算法等网格算法以实现复合指标计算;景观指数指标计算接口用于调用景观类型丰富度计算方法、景观破碎度计算方法、香农多样性计算方法等网格算法;距离指标计算接口用于调用网格距离算法、设施服务范围算法、网格服务效能算法等网格算法。

[0163] 作为非限制性示例,所述诊断型应用实现如下:

构建自定义规则的推理引擎,组件包括规则解析器,图谱匹配器以及异常验证器,实现“区域-网格-事件”三级异常诊断,并对异常原因开展归因研究。

[0164] 示范性地,基于规则推理算法,实现业务链路异常检测。

[0165] 示范性地,基于社区发现算法,实现土地利用变化异常区域检测。

[0166] 示范性地,基于归因分析算法,实现异常现象发生原因的推理。

[0167] 作为非限制性示例,所述预测型应用实现如下:

调用时空推理算法,实现事件预测、属性预测和空间分布预测,基于预测结果开展风险预警和响应建议。

[0168] 上述优选实施例中,提供了网格算法和图谱算法,可以进行复合指标计算、景观指

数计算、距离指标计算等,还能够实现规则推理、时空推理、因果推理和图推理等多种推理分析,为土地资源保护、土地规划等决策提供科学依据,支撑决策的科学性和合理性。

[0169] 在一些实施例中还提供一种计算机可读存储介质,所述存储介质上存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、至少一段程序、代码集或指令集由处理器加载并执行,使得所述处理器执行本申请前述实施例中所提供的方法的部分或全部步骤。

[0170] 可以理解,所述存储介质可以是瞬时性的,也可以是非瞬时性的。示范性地,所述存储介质包括但不限于U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机访问存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0171] 示范性地,所述处理器可以为中央处理器(Central Processing Unit,CPU)、微处理器(Microprocessor Unit,MPU)、数字信号处理器(Digital Signal Processor,DSP)或现场可编程门阵列(Field Programmable Gate Array,FPGA)等。

[0172] 示范性地,所述只读存储器包括但不限于MASK ROM、PROM、EPROM、EEPROM、Flash等。

[0173] 示范性地,所述随机访问存储器包括但不限于DRAM、SRAM、SDRAM、DDR SDRAM等。

[0174] 在一些示例中提供一种计算机程序产品,具体可以通过硬件、软件或其结合的方式实现。作为非限制性示例,所述计算机程序产品可以体现为所述存储介质,还可以体现为软件产品,例如SDK(Software Development Kit,软件开发包)等。

[0175] 作为非限制性示例,提供一种计算机程序产品,该计算机程序产品包括计算机程序或计算机可执行指令,该计算机程序或计算机可执行指令存储在计算机可读存储介质中。电子设备的处理器从计算机可读存储介质读取该计算机程序或者计算机可执行指令,处理器执行该计算机可执行指令,使得该电子设备执行本申请实施例所述方法的部分或全部步骤。

[0176] 在一些示例中提供一种计算机程序,包括计算机可读代码,在所述计算机可读代码在计算机设备中运行的情况下,所述计算机设备中的处理器执行用于实现所述方法中的部分或全部步骤。

[0177] 本实施例还提出一种电子设备,包括存储器和处理器,所述存储器存储有至少一条指令、至少一段程序、代码集或指令集,所述处理器执行所述至少一条指令、至少一段程序、代码集或指令集时实现如前述实施例中所提供的方法的部分或全部步骤。

[0178] 在一些示例中提供一种所述电子设备的硬件实体,包括:处理器、存储器和通信接口;其中,所述处理器通常控制所述电子设备的总体操作;所述通信接口用于使所述电子设备通过网络与其他终端或服务器通信;所述存储器配置为存储由处理器可执行的指令和应用,还可以缓存待处理器以及电子设备中各模块待处理或已经处理的数据(包括但不限于图像数据、音频数据、语音通信数据和视频通信数据),可以通过闪存(FLASH)、可擦写可编程只读存储器(EPROM,Erasable Programmable Read-Only Memory)、带电可擦可编程只读存储器(EEPROM,Electrically Erasable Programmable Read-Only Memory)或随机访问存储器(RAM,Random Access Memory)实现。

[0179] 处理器可包括一个或多个处理元件。因此,处理器可包括被配置为执行处理器的功能的一个或多个集成电路(IC)。此外,每个集成电路可包括被配置为执行处理器的功能

的电路(例如,第一电路、第二电路,以及其他电路等)。

[0180] 进一步地,处理器、通信接口和存储器之间可以通过总线进行数据传输,总线可以包括任意数量的互联的总线和桥,总线将一个或多个处理器和存储器的各种电路连接在一起。

[0181] 相同或相似的标号对应相同或相似的部件;

附图中描述位置关系的用语仅用于示例性说明,不能理解为对本申请的限制;

需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0182] 在不同的具体实施中,本申请所述的方法或系统可以在软件、硬件或它们的组合中实现。此外,可改变方法的步骤的顺序,并且可添加、重新排序、组合、省略、修改各种元素等。

[0183] 显然,本申请的上述实施例仅仅是为清楚地说明本申请所作的举例,而并非是对本申请的实施方式的限定,不用于限定本申请。对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其它不同形式的变化或变动,各分立的结构/功能模块或单元可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或更多个模块集成形成一个独立的部分,分立部件的结构和功能可被实现为组合的结构或部件。这里无需也无法对所有的实施方式予以穷举。凡在本申请的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本申请权利要求的保护范围之内。

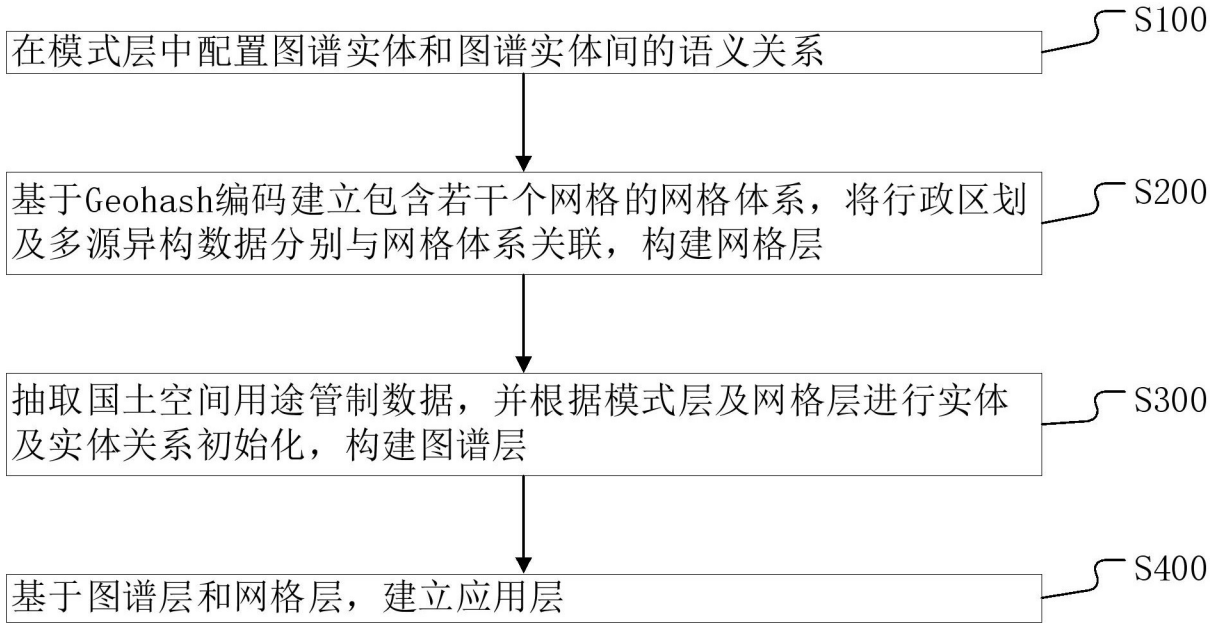


图 1

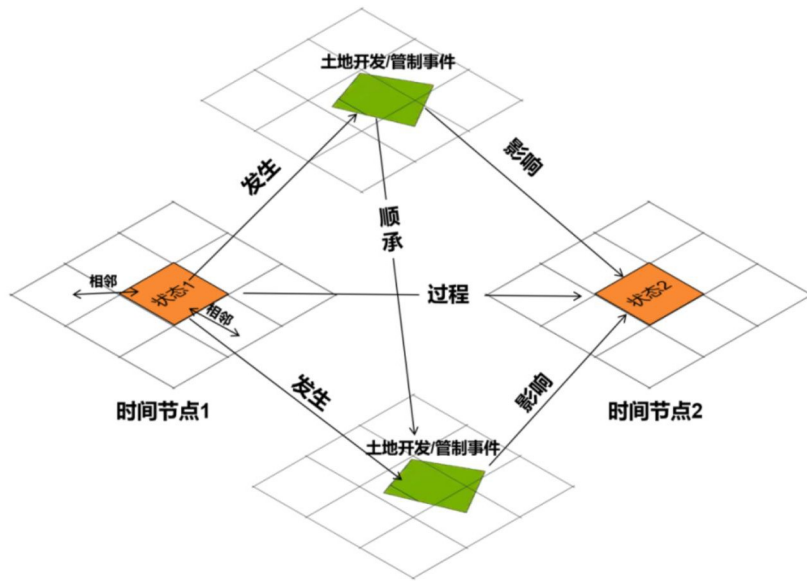


图 2

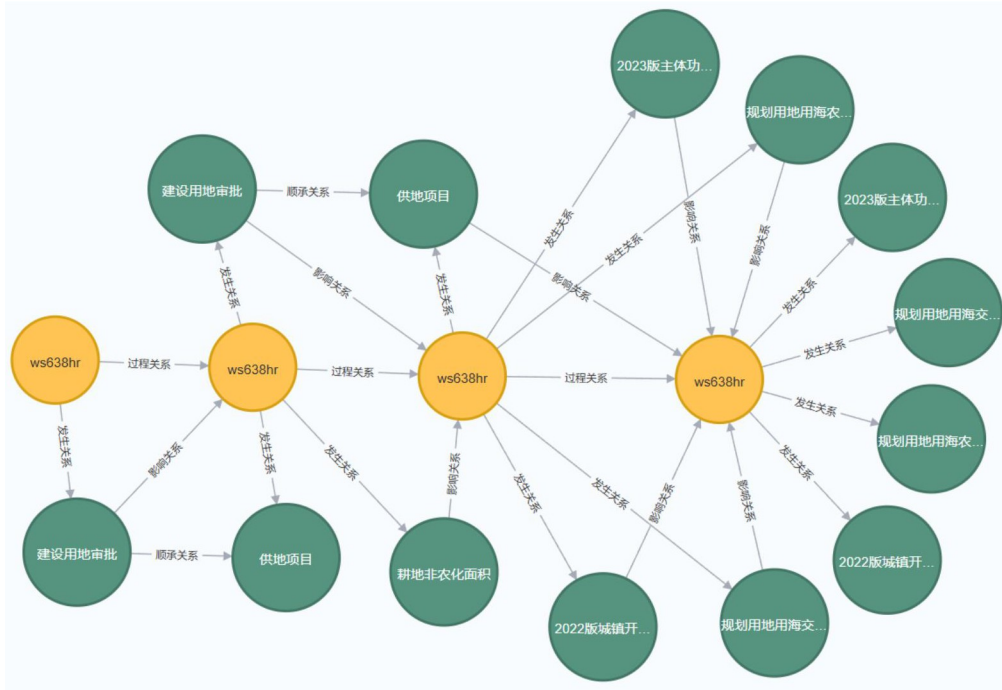


图 3

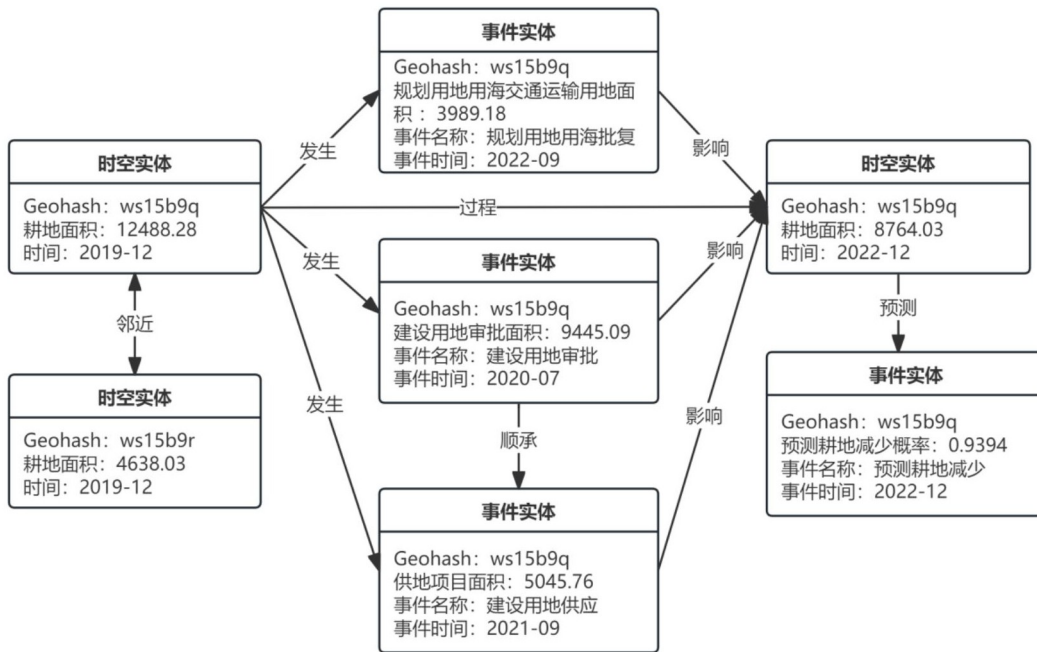


图 4

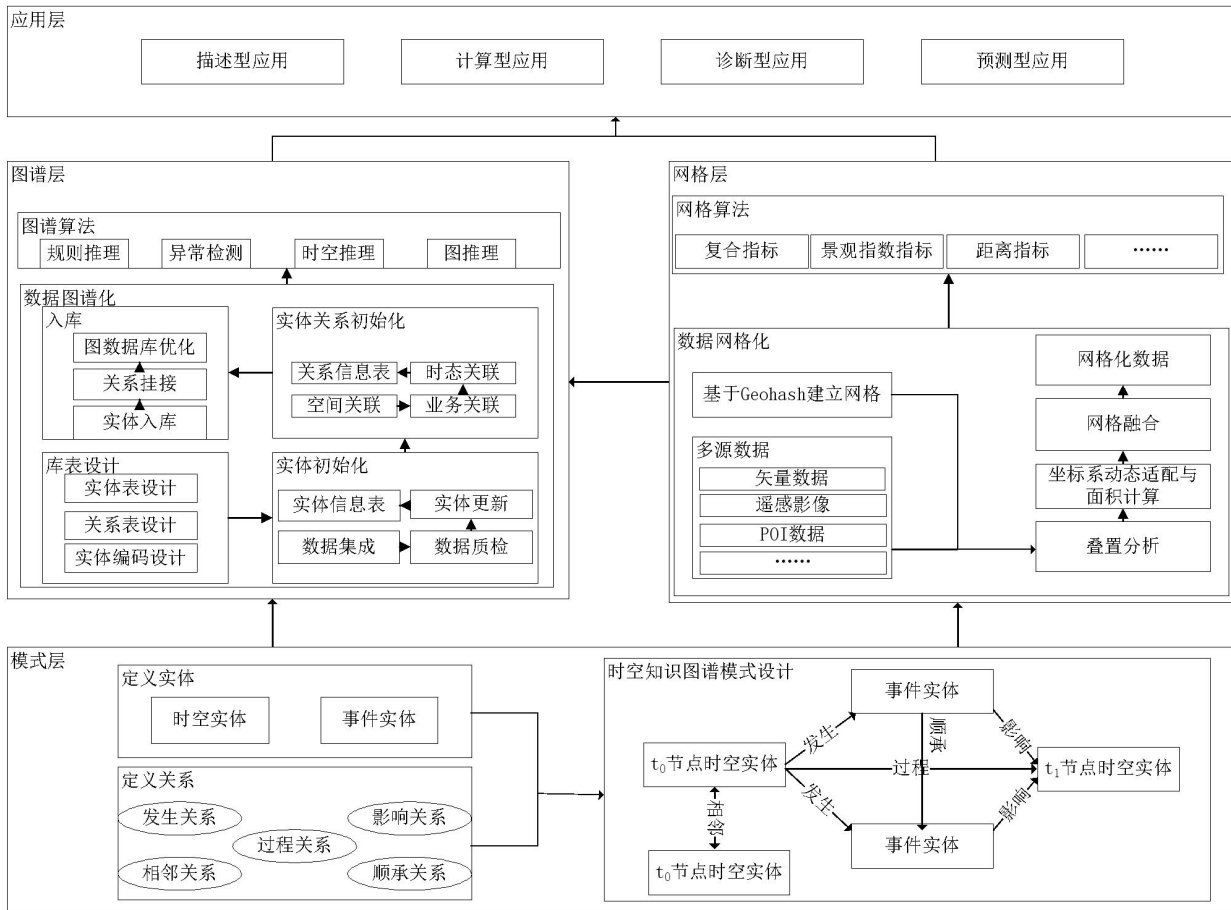


图 5