



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： 基于地理案例的城市规划信息组织
作者： 李锐，王璟琦，王顺利，李江，吴华意
DOI： 10.13203/j.whugis20220395
收稿日期： 2023-05-04
网络首发日期： 2023-06-02
引用格式： 李锐，王璟琦，王顺利，李江，吴华意. 基于地理案例的城市规划信息组织 [J/OL]. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20220395>



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.whugis20220395

引用格式：

李锐, 王璟琦, 王顺利, 等. 基于地理案例的城市规划信息组织[J]. 武汉大学学报(信息科学版), 2023, DOI: 10.13203/j.whugis20220395 (LI Rui, WANG Jingqi, WANG Shunli, et al. Information Organization in Urban Planning Based on Geographical Case[J]. *Geomatics and Information Science of Wuhan University*, 2023, DOI: 10.13203/j.whugis20220395)

基于地理案例的城市规划信息组织

李锐^{1,*}, 王璟琦¹, 王顺利¹, 李江², 吴华意¹

1 武汉大学 测绘遥感信息工程国家重点实验室, 湖北 武汉, 430079;

2 湖北省自然资源厅 信息中心, 湖北 武汉, 430071)

摘要：随着我国城镇化进程的快速推进，持续扩大的人口规模与有限空间资源之间的矛盾为城市发展和规划建设带来了诸多挑战，深刻影响着民众对城市生活空间的观点情感。如何充分利用地理数据和舆情大数据资源，对城市规划中多维性和复杂性的地理信息与网络舆情信息进行有序组织成为研究的难点。本文基于地理案例思想提出城市规划案例的概念，对城市规划中的地理信息和舆情信息进行有序组织。首先构建了包含<案例问题，案例属性，案例结果>的城市规划案例描述模型；基于此，对城市规划案例的问题、属性及结果进行量化表达，建立了城市规划案例表达框架；最后，以具体的规划项目为例，对城市规划案例描述及表达方法进行实例说明。该方法的提出，实现了城市规划的地理信息与网络舆情信息抽象化描述与量化表达，减少了城市规划案例特征组成的冗余，同时为城市规划案例相似度计算提供理论支持。

关键词：城市规划案例；网络舆情；地理案例；城市规划信息组织

中图分类号：TU984

文献标志码：A

收稿日期：2023-05-04

Information Organization in Urban Planning Based on Geographical Case

LI Rui^{1,*}, WANG Jingqi¹, WANG Shunli¹, LI Jiang², WU Huayi¹

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China;

2 Information Center, Department of Natural Resources of Hubei Province, Wuhan 430071, China

Abstract: Objectives: With the rapid urbanization in China, the contradiction between the ever-expanding population size and the limited space resources has brought many challenges to urban and rural planning and construction management, and also affected the people's emotional attitude towards urban space. How to make full use of geographic data and public opinion big data, to organize multidimensional and complex geographic information and online public opinion information in urban planning has become a difficult issue in the research of online public opinion. Based on the idea of geographical case, this paper proposes the concept of urban planning case and organizes geographical information and public opinion information in urban planning in an orderly way. **Methods:** The urban planning case description model is constructed according to the triplet form of "case problem, case attribute, and case result". On this basis, case problem, case attribute and case result of urban planning case are quantitatively expressed, and the expression framework of urban planning case is constructed. **Results:** Finally, a specific planning project is taken as an example, to illustrate the description and expression method of urban

收稿日期：2023-05-04

基金项目：国家自然科学基金 (U20A2091, 41930107)。

第一作者：李锐，博士，教授，主要研究方向为时空计算与大数据挖掘、语义计算、网络空间行为分析及分布式实时系统。ruili@whu.edu.cn

planning case. **Conclusions:** The proposed method realizes the abstract description and quantitative expression of geographic information and online public opinion information in urban planning, reduces the redundancy of the composition of the urban planning network public opinion, and provides theoretical support for the similarity calculation of urban planning cases.

Key words: urban planning case; network public opinion; geographical case; organization of urban planning information

在全球快速城市化的背景下,提高城市生活质量正成为城市规划中日益重要的问题^[1]。持续增长的人口与不断减少的可用空间资源之间的冲突日趋明显,影响着人们的生活质量和社会幸福感^[2]。公众参与到城市规划的愿望越来越强烈,许多学者和从业者已经注意到公民参与城市规划实践的广泛多样性^[3,4,5]。公众借助社交网络、政府网站等网络公开平台参与城市规划,形成了大量的城市规划网络舆情,影响着城市规划的合理编制与顺利实施^[6,7]。如何充分利用丰富的地理数据和海量的舆情数据资源,增强对舆情数据的语义理解能力,度量城市规划地理特征及舆情特征,成为亟待研究的问题。

本文中城市规划信息包括城市规划中的地理信息和舆情信息,城市规划信息组织是对城市规划中的地理信息和舆情信息进行描述、表达和提取,科学准确地反映规划项目的时空属性及舆情特征。在城市规划的地理信息组织方面,面向对象时态GIS^[8]、面向决策支持的数据模型^[9]等方法被用来表达城市规划中的时间、空间和属性信息。Frederico等^[10]首次提出了本体论驱动的地理信息系统,结合语义、空间和属性的表达框架逐步形成^[11]。闫国年等^[12]从地理学视角提出地理信息六要素表达,提升了地理信息的组织、管理、表达和分析能力。随着地理信息理论的发展以及GIS发展的时代需求,越来越多的研究不只关注地理实体自身内部的属性,转而重视多元空间内实体之间的关联^[13,14]。在城市规划的舆情信息组织方面,王卷乐等^[15]基于疫情的相关微博数据,从数量、空间、时间、内容等方面研究区域尺度的舆情分布特征。董坚峰^[16]以在公共危机预警中起关键作用的话题、事件、网民个体、网民群体等舆情要素作为分析对象,分别从话题发现、事件识别、网民个体分析、群体行为分析四个方面探讨网络舆情分析的主要模式。Zhou Zhipeng等^[17]提出了结合生命周期和区域特征的舆情分析框架,用于大数据背景下大型基础设施的主题建模和情感分析,并对港珠澳大桥建设项目的微博舆情

进行分析。当前网络舆情研究主要关注于舆情信息的提取,对多维性和复杂性的城市规划信息进行有序组织研究还相对薄弱,难以解决城市规划网络舆情组成的冗余和实现城市规划案例的抽象化描述与量化表达。

本文以案例推理为理论依据,基于“地理案例”表达方法来组织城市规划的地理环境特征和网络舆情特征,提出城市规划案例描述模型及表达框架,实现城市规划案例信息的抽象化描述与量化表达,可为规划项目管理与制定提供决策支持。

1 城市规划案例描述与表达

地理案例推理是一种利用案例推理来研究分析地理数据与地理问题的方法。其研究对象是地理案例,即对特定时空中发生的地理现象的一种解析与重构形式^[18],通常以三元组<案例问题,案例属性,案例结果>进行描述。其中,“问题”指拟研究处理的地理问题,“属性”包含了地理案例自身的时空特征及所处地理环境的关系特征,“结果”对应地理案例的状态信息。在地理信息系统领域中,地理案例推理广泛用于地理过程的描述^[19]、空间关联分析^[20]等。

1.1 城市规划案例的描述模型

城市规划案例是基于地理案例对城市规划项目相关的地理环境信息和网络舆情信息的抽象化描述,是传统地理案例在城市规划领域的拓宽。它不仅包括地理环境信息,还含有丰富的网络舆情信息。随着物联网的出现和互联网的日益普及,公众借助网络公开平台参与城市规划,形成了大量的城市规划网络舆情,影响着城市规划的合理编制与顺利实施^[21,22]。充分考虑舆情中公众的情感态度,有助于城市规划的合理编制和顺利实施^[23]。

本文基于<案例问题,案例属性,案例结果>对城市规划的地理环境信息与网络舆情信息进行

组织，描述城市规划案例中的网络舆情信息与地理信息之间的关系。城市规划案例的描述模型如图1所示。

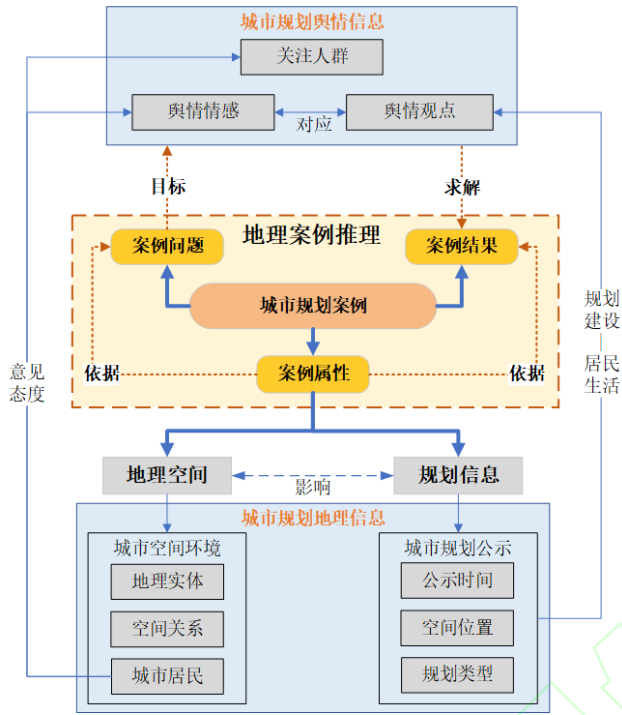


图1 城市规划案例描述模型

Fig.1 The descriptive model of Urban Planning Case

(1) 案例问题

本文定义案例问题为城市规划案例的多维度舆情特征预测问题，旨在获取未来某一城市规划案例的关注人群类别、多维度观点特征和多粒度情感特征。关注人群指关心规划项目编制与实施的居

民；观点类别是居民主张所聚焦的规划管理的各个方面，例如居住条件、安全隐患等；情感倾向是指城市居民个体或群体对于规划项目所表现出的情感态度或偏好，代表着居民对于规划项目编制的满意程度。

(2) 案例属性

案例属性指城市规划案例的地理信息，包括规划项目的规划信息与所处地理空间的环境特征。规划信息包含案例的规划时间、规划地点、规划类别，环境特征不仅包含空间地理实体的位置及功能类别，还包括生活在该空间中的城市居民的属性特征，即通过公开途径表达对案例观点和看法的公众。

(3) 案例结果

案例结果指的是案例问题所对应的结果和状态，即网络舆情情感和观点特征。在城市规划案例中，网络舆情信息可分为粗粒度和细粒度两类。粗粒度舆情特征指案例所涉及的人群类别和总体情感倾向分布，而细粒度舆情特征则是案例中观点的分类以及这些类别观点所对应的情感倾向。

1.2 城市规划案例的表达框架

本文顾及城市规划中的地理信息和舆情信息提出了城市规划案例的表达框架，实现了地理环境特征和网络舆情特征的量化表达，表达框架如图2所示。

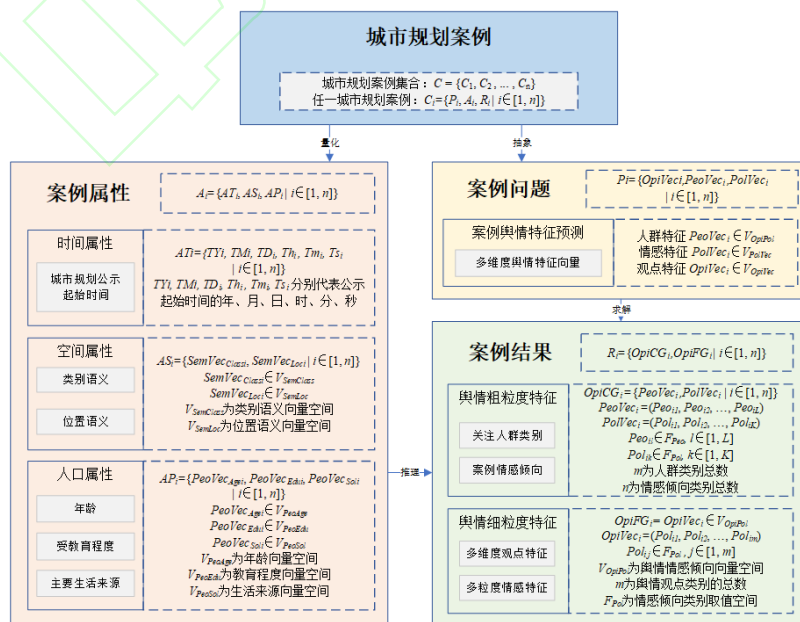


图2 城市规划案例表达框架

Fig.2 The expression framework of Urban Planning Case

本文通过构建<案例问题, 案例属性, 案例结果>三元组, 实现了城市规划案例的向量化表达。其中, 案例问题定义了人群特征、情感特征、观点特征等舆情特征; 案例属性量化了时间属性、空间属性和人口属性等属性特征; 案例结果明确了关注人群类别、案例情感倾向、多维度观点特征和多维度情感特征的量化表达。基于这一表达框架, 可以实现城市规划中地理信息与网络舆情信息的有序组织与定量化度量。

2 城市规划案例信息提取方法

2.1 案例属性: 地理环境特征

城市规划案例在空间上与邻近环境密切相关, 案例的地理环境特征受城市中地理实体的功能属性及空间关联影响。因而, 在表达城市规划案例的案例属性时, 既要顾及规划项目的时间、位置等信息, 也需要探究案例所处环境的功能特征与结构特征。空间位置是城市规划案例的重要要素^[24], 需要结合案例的位置特征与周边环境的位置特征来进行表征。考虑到地理环境的特征类型, 将案例属性分为时间属性、空间属性和人口属性三类。

2.1.1 时间属性

城市规划案例的时间属性通常以规划项目开始进行社会公示的时间来表示。案例的时间属性既代表规划项目社会公示的起点时刻, 也表明引起关于规划项目网络舆情的起始时间。定义规划案例 C_i 的时间属性为 $AT_i = \{TY_i, TM_i, TD_i, Th_i, Tm_i, Ts_i | i \in [1, n]\}$, 其中 $TY_i, TM_i, TD_i, Th_i, Tm_i, Ts_i$ 分别代表时间信息的年、月、日、时、分、秒, 对于低精度的部分用0表示。

2.1.2 空间属性

空间语义是指在某个领域内, 对时空数据进行具体的解释和表达^[25]。本文定义城市规划案例的空间语义包括两个方面: 案例类别语义和案例位置语义。

(1) 案例类别语义

案例类别语义反映了规划项目所处区域的用地功能特征, 即规划项目所承担的城市建设功能。为实现空间类别语义的量化表达, 本文结合规划用地类别对城市规划案例进行分类, 并基于规划领域语料库 $Corpus_{Case}$ 形成规划领域词向量, 从而实现

案例类别词汇到语义空间的映射。本文定义案例类别语义的量化表达形式为 $SemVec_{Class_i} = (W_{Class_{i1}}, W_{Class_{i2}}, \dots, W_{Class_{iM}})$, 通过表示学习方法获取案例类别词汇对应的语义向量。具体流程如图3所示。

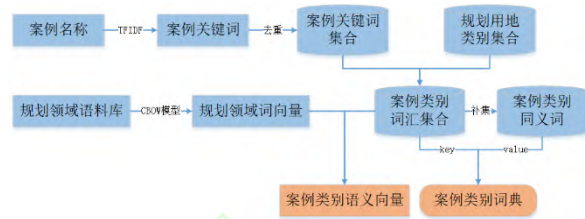


图3 城市规划案例的类别语义计算

Fig.3 The calculation process of category semantics of Urban Planning Case

针对城市规划案例集合 $C = \{C_1, C_2, \dots, C_n\}$, 其中 n 为案例总数, 其对应的案例名称集合为 $Name = \{Name_1, Name_2, \dots, Name_n\}$ 。任意一个案例 C_i 的名称 $Name_i$ 均可用中文分词后的词汇序列来表示, 即 $Name_i = \{word_1, word_2, \dots, word_m\}$, 其中 m 为名称 $Name_i$ 中的词汇总数。对于 $Name_i$ 中任意词汇 $word_j, j \in [1, m]$, 其权重 w_{ij} 计算方法如公式(1)所示。其中, $count(i, j)$ 表示词汇 $word_j$ 在案例名称 $Name_i$ 中出现的次数, $count(j, n)$ 则表示案例名称集合 $Name$ 中包含词汇 $word_j$ 的案例名称数量。

$$w_{ij} = \frac{count(i, j)}{m} \cdot \log \frac{n}{count(j, n)} \quad (1)$$

针对每个案例名称 $Name_i$, 选取权重 w_{ij} 最大的 $\frac{m}{2}$ 个词汇作为案例关键词 $Kwds^{Name_i}$, 同时将所有案例的关键词合并为一个案例关键词集合 $Kwds^{Case}$ 。在此基础上, 依据国家标准《城市用地分类与规划建设用地标准》的内容范围^[26], 获取其中的用地类别名称并建立规划用地类别集合 $Class^{Land}$ 。根据公式(2), 案例类别集合定义为案例关键词集合与规划用地类别集合的交集。 $Kwds^{Case}$ 中与 $Class^{Case_i}$ 共同出现在同一案例名称中的词汇称为案例类别同义词, 对应的所有同义词可以表示为集合 Syn^{Case_i} , 构成键值对 $\langle Class^{Case_i}, Syn^{Case_i} \rangle$ 。通过计算所有案例类别的键值对, 形成案例类别词典 $ClassDict^{Case}$ 。

$$Class^{Case} = Kwds^{Case} \cap Class^{Land} \quad (2)$$

以城市规划相关的新闻报道为基础, 构建城市规划领域的文本语料库 $Corpus_{Case}$, 并将其作为训练数据以获得案例类别语义向量。对于案例类别集合为 $Class^{Case}$ 中的任一案例类别 cls_{ci} , 采用滑动窗口的方法获得对应的位置上下文, 记为

$Context(cls_{ci}) = cls_{ci-m}^{ci+m}$ ，其中 m 表示窗口大小 cls_{ci-m}^{ci+m} 表示以 cls_{ci} 为中心的 $2m$ 类别词汇序列 $\{cls_{ci-m}, \dots, cls_{ci-1}, cls_{ci+1}, \dots, cls_{ci+m}\}$ 。利用模型投影层来获取输入层的上下文向量之和，进而得到案例类别词汇 cls_{ci} 对应的语义向量 v_{ci} 。

(2) 案例位置语义

相较于传统以空间坐标表示位置的方式，案例位置语义可以准确描述案例的空间位置，并能定量度量空间分布特征。本文提出一种名为 Loc2Vec (Location to Vector) 的案例位置语义方法，顾及规划项目与周边 POI 的空间距离，来计算并获得案例位置语义向量，定义案例位置语义的量化表达形式为 $SemVec_{Loc_i} = (w_{Loc_{i1}}, w_{Loc_{i2}}, \dots, w_{Loc_{iN}})$ 。Loc2Vec 方法流程如图 4 所示。

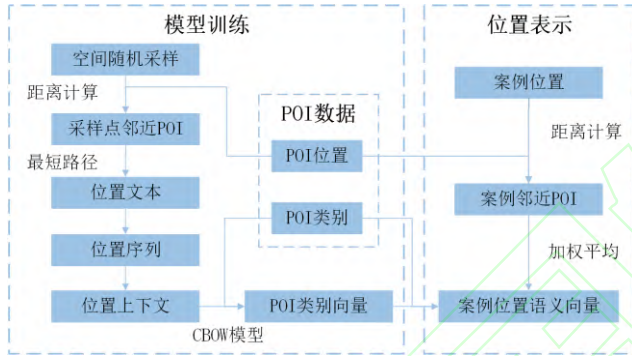


图4 Loc2Vec方法流程

Fig.4 The process of the Loc2Vec method

案例位置语义是通过周边 POI 点的类别语义进行加权平均而得到，权重由 POI 点与规划项目位置间的距离倒数计算而得。针对城市规划案例 C_i ，定义其周边 POI 集合为 $NPP(C_i) = \{p_1, p_2, \dots, p_k\}$ ，其中 p_j 表示 POI 集合中得第 j 个元素， k 为集合中 POI 点的总数。案例 C_i 对应的位置语义向量 $SemVec_{Loc_i} = (w_{C_{i,1}}, w_{C_{i,2}}, \dots, w_{C_{i,r}})$ 中第 s 维 $w_{C_{i,s}}$ 的权重计算方法如公式 (3) 所示。其中 $dist(p_j, C_i)$ 表示点 p_j 与案例 C_i 之间的距离， $w_{p_{j,s}}$ 为 p_j 对应的类别语义向量 $(w_{p_{j,1}}, w_{p_{j,2}}, \dots, w_{p_{j,r}})$ 的第 s 维。

$$w_{C_{i,s}} = \frac{\sum_{j=1}^k \left(\frac{1}{dist(p_j, C_i)} \times w_{p_{j,s}} \right)}{\sum_{j=1}^k \frac{1}{dist(p_j, C_i)}} \quad (3)$$

对于任意两个案例位置语义向量 $v_i = (w_{i1}, w_{i2}, \dots, w_{in})$ 、 $v_j = (w_{j1}, w_{j2}, \dots, w_{jn})$ ，本文采用余弦距离 (CosDist) 和相关系数 (Corrcoef) 作为向量相似度度量指标。余弦距离的计算方法如公式 (4) 所示，相关系数的计算方法如公式 (5) 所示。

$$CosDist(v_i, v_j) = 1 - \frac{\sum_{l=1}^n (w_{il} \times w_{jl})}{\sqrt{\sum_{l=1}^n (w_{il}^2) \times \sum_{l=1}^n (w_{jl}^2)}} \quad (4)$$

$$Corrcoef(v_i, v_j) = \frac{\sum_{k=1}^n (w_{ik} - \bar{w}_i) (w_{jk} - \bar{w}_j)}{\sqrt{\sum_{k=1}^n (w_{ik} - \bar{w}_i)^2 \sum_{k=1}^n (w_{jk} - \bar{w}_j)^2}} \quad (5)$$

2.1.3 人口属性

城市规划案例的人口属性指针对城市规划项目表达观点和情感态度的城市居民的属性，这一属性描述了城市居民在规划案例所涉及的空间环境中的不同特征。不同城市居民因在年龄、学历、生活环境等方面存在差异，从而影响其对于城市规划案例的关注方面和情感倾向 [27, 28]。为实现人口属性的量化表达，本文综合考虑了城市居民不同维度的特征，包括年龄、受教育程度、主要生活来源，并基于案例所在的空间环境进行研究。

(1) 人口年龄

将人口年龄划分为 k 个等间隔的区间，定义案例 C_i 的人口年龄特征向量为 $PeoVec_{Age_i} = (w_{Age_{i1}}, w_{Age_{i2}}, \dots, w_{Age_{il}}, \dots, w_{Age_{ik}})$ 。其中，该向量的维数为 k ，每一维度的值 $w_{Age_{il}}$ 表示第 l 年龄区间内的人数占总人口数的比例。

(2) 受教育程度

根据城市居民学历的差异，将其受教育程度划分为 s 类，定义案例 C_i 的受教育程度特征向量为 $PeoVec_{Edu_i} = (w_{Edu_{i1}}, w_{Edu_{i2}}, \dots, w_{Edu_{il}}, \dots, w_{Edu_{is}})$ 。该向量的维数为 s ，每一维度的值 $w_{Edu_{il}}$ 表示第 l 类受教育程度的人数占总人口数的比例。

(3) 主要生活来源

根据生活来源的不同，将其划分为 q 类，定义案例 C_i 的主要生活来源特征向量为 $PeoVec_{Soli_i} = (w_{Soli_{i1}}, w_{Soli_{i2}}, \dots, w_{Soli_{il}}, \dots, w_{Soli_{iq}})$ 。该向量的维数为 q ，每一维度的值 $w_{Soli_{il}}$ 表示第 l 类主要生活来源的人数占总人口数的比例。

2.2 案例问题和案例结果：网络舆情特征

网络舆情特征是案例问题和案例结果求解的关键，本文考虑网络舆情的粗粒度和细粒度特征，实现了网络舆情特征的量化表达，并分别采用词汇级和篇章级的方法提取文本中不同级别的语义信息。案例的粗粒度舆情特征包括关注人群类别和案例情感倾向等宏观属性，细粒度舆情特征包括多维度观点特征及与观点对应的多维度情感特征。案例

舆情的多粒度特征提取方法框架如图 5 所示。

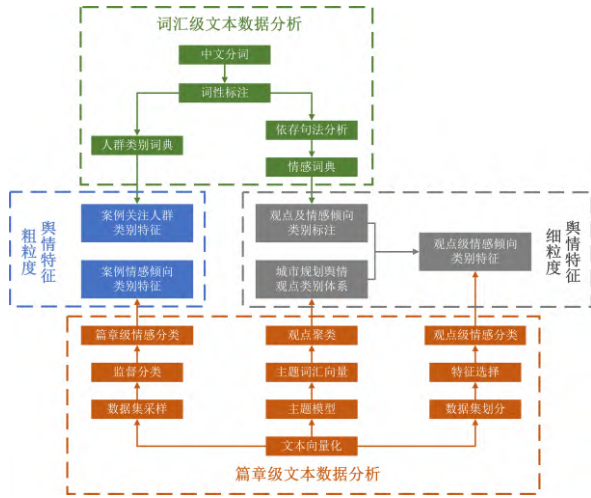


图5 案例舆情的多粒度特征提取方法框架

Fig.5 The framework of multi-granularity features' extraction method for case public opinion

2.2.1 粗粒度舆情特征

规划案例的粗粒度舆情特征主要涵盖关注人群类别和案例情感倾向两个方面。关注人群类别表示根据相关的标准或特征划分得到的城市居民类型，关注人群类别的提取有助于了解不同类别居民对规划项目的需求和期望。案例情感倾向反映城市居民对规划项目的群体情感，情感倾向可分为正面、负面和中性三种类型。

(1) 关注人群类别

出于对公民隐私的保护，城市规划中的舆情数据通常具有匿名性，城市居民的诸多个人信息无法直接获取。针对这一问题，本文提出一种从舆情文本中进行关注人群提取的方法，具体流程如图 6 所示。首先，采用用户词典和词性标注方法从舆情文本中提取关注人群类别词汇，形成关注人群类别词典 $PeoT^L = \{t_1, t_2, \dots, t_L\}$ ， L 为类别总数；其次，统计案例 C_i 的网络舆情中涉及的不同类别人群 t_l ，建立案例 C_i 的关注人群类别集合 Peo_i ；最后，基于公式 (6) 将集合 Peo_i 映射为 L 维向量，得到案例 C_i 的关注人群类别向量 $PeoVec_i = (Peo_{i1}, Peo_{i2}, \dots, Peo_{iL})$ ，若案例 C_i 中包含第 l 类人群则 $Peo_{il} = 1$ ，反之 $Peo_{il} = 0$ 。

$$Peo_{il} = \begin{cases} 0, & t_l \notin Peo_i \\ 1, & t_l \in Peo_i \end{cases} \quad (6)$$



图6 案例舆情的关注人群提取流程

Fig.6 The extraction process of the concerned crowd of case public opinion

(2) 案例情感倾向

本文采用多类别分类方法来获取案例情感倾向，记案例 C_i 中的所有反馈意见文本为 $D_i = \{d_{i1}, d_{i2}, \dots, d_{im}\}$ ，情感倾向类别的集合为 $PolT^K = \{t_1, t_2, \dots, t_K\}$ ，按图 7 所示的情感倾向提取流程可获得对应的情感倾向序列 $Pol_i = \{p_{i1}, p_{i2}, \dots, p_{im}\}$ 。用 K 维向量 $PolVec_i = (Pol_{i1}, Pol_{i2}, \dots, Pol_{iK})$ 来表示 C_i 的案例情感倾向， Pol_{ik} 的计算方法如公式 (7)。其中 $count(Pol_i, t_k)$ 表示列表 Pol_i 中值为 t_k 的元素数目， m 为列表总长度。

$$Pol_{ik} = \frac{count(Pol_i, t_k)}{m} \quad (7)$$

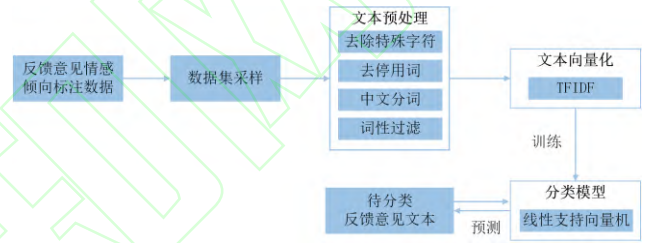


图7 案例情感倾向的提取流程

Fig.7 The extraction process of emotional tendency of the case

2.2.2 细粒度舆情特征

城市规划案例的细粒度舆情特征包括多维度观点特征和多维度情感特征，可以更全面地描述城市规划舆情的的情感倾向，并且可以在多个观点类别上进行细致的情感分析。

(1) 多维度观点特征

使用隐含狄利克雷分布 (Latent Dirichlet Allocation, LDA) 模型获得 K 类观点的集合 $\varphi = \{\varphi_k | k \in \{1, 2, \dots, K\}\}$ ，其中每类观点 φ_k 包含 Q 个词汇向量，即 $\varphi_k = (\varphi_{k1}, \dots, \varphi_{kq}, \dots, \varphi_{kQ})$ 。为了将这些词汇向量聚类成多层次的观点聚类簇，首先将每类观点定义为一个聚类簇，按照公式 (8) 计算两个聚类簇间的最短词汇距离 $dist(\varphi_{im}, \varphi_{jn})$ 作为两个聚类簇之间的距离。然后，基于最小距离准则对初始聚类簇进行合并，直到所有词汇向量都被合并到一个观点聚类簇中。

$$d(u_i, u_j) = \min(dist(\varphi_{im}, \varphi_{jn})) \quad (8)$$

(2) 多维度情感特征

城市规划案例的多维度情感特征与案例的多维度观点特征相对应，将其记为 $OpiVec_i = (Pol_{i1}, Pol_{i2}, \dots, Pol_{iP})$ ，这一特征是各个文本中观点

所对应的情感倾向的汇总。对于任意一条反馈意见文本 $d_{ij} \in D_i$ ，其多维观点对应的情感倾向为 $OpiVec_{d_{ij}} = (Pol_{ij1}, Pol_{ij2}, \dots, Pol_{ijp}, \dots, Pol_{ijJ})$ ，其中 Pol_{ijp} 代表该文本中对于某一观点的情感倾向。在获得所有文本的多维情感倾向的基础上，根据公式 (9) 计算案例的情感倾向 Pol_{ip} 。其中 Pol_{ijp}^s 代表文本 d_{ij} 的第 p 维观点对应的情感倾向类别取值，即统计案例的 J 个反馈意见文本中关于第 p 维观点的情感类别分布情况，数目最多的情感类别代表第 p 维观点的情感类别 Pol_{ip} 。

$$Pol_{ip} = \underset{s}{\operatorname{argmax}} \operatorname{count}(D_i, Pol_{ijp}^s) \quad (9)$$

3 实验与分析

3.1 研究数据

本文使用的数据来源包括规划项目的批前公示数据与公众反馈意见、POI数据和人口数据。

(1) 规划项目公示数据和公众反馈意见

城市规划项目公示及相应的公众反馈意见数据来自某市规划和自然资源局网站。公众可以在该网站自由浏览和下载规划公示数据，包括城市规划项目的项目编号、名称、位置等详细信息。规划公示的公众意见也通过该网站征集，每条反馈意见数据记录包括评论时间、评论的项目编号、评论内容等。规划公示和反馈意见可以通过项目编号关联，一个规划项目对应多条反馈意见。数据的时间范围为2007年11月22日至2019年9月3日，含有3,434个规划项目、102,825条反馈意见。

(2) POI数据

POI数据通过高德地图的API获取，包括编号ID、名称、类别、坐标等字段。共获得POI数据893,924条，经过处理转换为WGS84坐标。POI数据示例如表1所示。

表1 POI数据示例

Tab. 1 Example of POI data

字段名称	字段内容
编号	q#20180707#992ceddadaa424985ce600ddd7acada
名称	**酒店(**东路)
类别	住宿服务,宾馆酒店,五星级宾馆
行政区编号	440106
坐标	113.373423,23.124661

(3) 人口数据

人口数据来自某市统计局官方网站公布的人口普查数据，主要包括人口年龄、教育程度和主要生活来源三类信息。根据本文2.1.3节中定义的人口属性，将年龄划分为1岁以下、1-4岁、5-9岁、...、95-99岁、100岁及以上等22个区间，并使用22维的向量表示人口年龄 (AgeVec)；受教育程度被划分为未上过学、小学、初中、高中、大学专科、大学本科、研究生等7个类别，用7维的向量表示受教育程度 (EduVec)；主要生活来源被划分为劳动收入、离退休金养老金、失业保险金、最低生活保障金、财产性收入、家庭其他成员供养、其他等7个类别，用7维的向量表示主要生活来源 (SolVec)。表2展示了部分案例的人口特征向量示例。

表2 人口特征向量示例

Tab. 2 Example of population eigenvectors

编号	AgeVec	EduVec	SolVec
1	0.014,0.044,0.046,...	0.013,0.142,0.319,0.35	0.460,0.293,0.001,0.00
2	0.089,0.096,...	7,0.104,0.057,0.005	9,0.002,0.202,0.028
3	0.007,0.026,0.035,...	0.018,0.254,0.528,0.14	0.715,0.053,0.001,0.00
	0.051, 0.079,...	6,0.036,0.014,0.001	4,0.005,0.201,0.020
	0.003,0.024,0.030,...	0.009,0.143,0.445,0.30	0.700,0.090,0.001,0.00
	0.036,0.042,...	0,0.067,0.030,0.002	3,0.021,0.164,0.018

3.2 城市规划案例总体特征分析

3.2.1 案例地理环境整体特征

为了探究研究数据的案例地理环境的整体特征，本文使用了两种距离度量方法：位置语义距离和空间坐标欧氏距离，分别计算了任意两个案例之间的距离。其中，案例间的位置语义距离根据公式 (4) 进行计算，并将计算结果作为横坐标，以案例间的空间坐标欧氏距离作为纵坐标，绘制了案例距离散点图 (如图8所示)。通过对案例距离散点图的分析，发现位置语义距离小的案例在空间坐标欧氏距离方面也较小，表明位置语义距离与空间坐标欧氏距离具有正相关关系。通过公式 (5) 计算二者的相关系数为0.799，验证了本文2.1.2节中提出的 Loc2Vec方法可以表示案例的空间位置特征，并实现了对案例空间位置属性的量化。

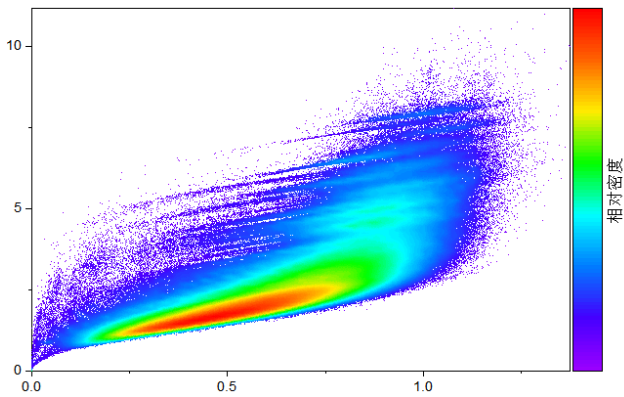


图8 案例位置语义向量余弦距离及地理坐标欧氏距离散点分布相对密度图

Fig.8 Scatter plot density of case location semantic vector cosine distance and Euclidean distance of geographic coordinates

3.2.2 案例网络舆情总体特征

(1) 案例关注人群特征

在城市规划案例中，通常会使用一些特定的词汇来描述关注人群的年龄、职业等特征，如“业主”、“病人”、“保安”、“老师”等，将这些词汇称为关注人群类别词汇。本文基于2.2.1节中定义的关注人群提取流程，对反馈意见文本进行处理可得到28个关注人群类别。对反馈意见文本提取结果进行汇总，获得规划案例的关注人群。随后，对这些类别涉及的案例数量进行统计，并通过图9展现。根据图8可知，涉及“业主”类别的案例最多，约占总数的57%，这表明城市规划对“业主”人群的影响最大。基于公式（6）计算案例的关注人群特征向量，其示例如表3所示。

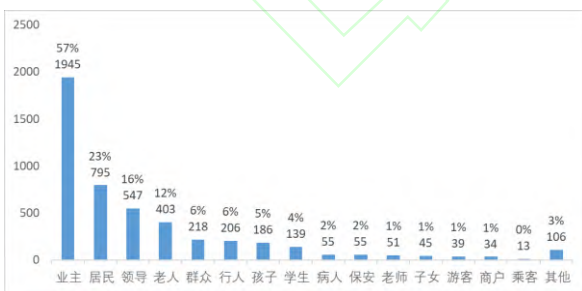


图9 关注人群相关的案例数目统计

Fig.9 Statistics on the number of cases related to the concerned population

表3 案例的关注人群特征向量示例

Tab. 3 Example of feature vector of concerned population of case public opinion

编号	案例名称	关注人群特征向量
----	------	----------

1	***医院地块调整	11111110100000 00010000000000
2	***小区地块更改	11101110100000 00000000000000
3	***广场天桥工程	11001111000100 10000000000000

(2) 观点类别体系

基于2.2.2节中提出的观点类别提取方法，获得了公众反馈意见的观点-词汇概率分布，并在表2中展示了其中部分观点。为了更好地理解每个观点的语义特征，按照词汇概率值进行降序排序，并选取概率最高的10个词汇作为观点的语义特征。根据这些词汇的含义定义每个观点的名称，并基于凝聚层次聚类方法获得多层次的观点聚类簇。图10展示聚类结果，其中横坐标表示不同观点聚类簇之间的距离，纵坐标为各初始观点的编号。

表4 观点-词汇概率分布

Tab.4 An example of viewpoint - vocabulary probability distribution

编号	观点名称	观点-词汇概率分布
15	规划公示	广州市 0.266349;海珠区 0.071829;工程 0.054272;优势 0.013960;规划 0.013035;响应 0.007965;公示 0.007597;影响 0.005798;展现 0.005318;建议 0.004614
37	资源保护	资源 0.150913;浪费 0.064571;取消 0.064152;保护 0.055309;专家 0.024032;规划 0.021574;中国 0.020650;矛盾 0.020226;保留 0.018307;地区 0.017784
26	用地调整	用地 0.233309;地块 0.207223;规划 0.187164;建设 0.047544;公示 0.045186;调整 0.038224;改为 0.032612;修改 0.014826;面积 0.013095;留用 0.007767
59	用地变更	大道 0.085626;用途 0.075550;性质 0.072501;密度 0.052542;完善 0.036683;规划 0.035293;投诉 0.033062;变为 0.023283;上访 0.023116;用地 0.016151
2	环境保护	公园 0.104943;宁静 0.069065;管理 0.055976;城市规划 0.042466;美丽 0.037764;居民 0.031640;设施 0.025574;敏感 0.024331;环境保护 0.016219;密集区 0.015071
38	环境垃圾	垃圾 0.524953;安宁 0.029802;理由 0.021624;影响 0.020941;减轻 0.020370;绿地面积 0.013706;反对 0.006314;金钱 0.006048;公摊 0.004269;杂物 0.004040
64	环境辐射	变电站 0.202809;居民 0.053027;反对 0.044152;距离 0.043186;高压 0.041842;居民区 0.038884;电磁辐射 0.035259;建设 0.033756;规划 0.025297;健康 0.023701

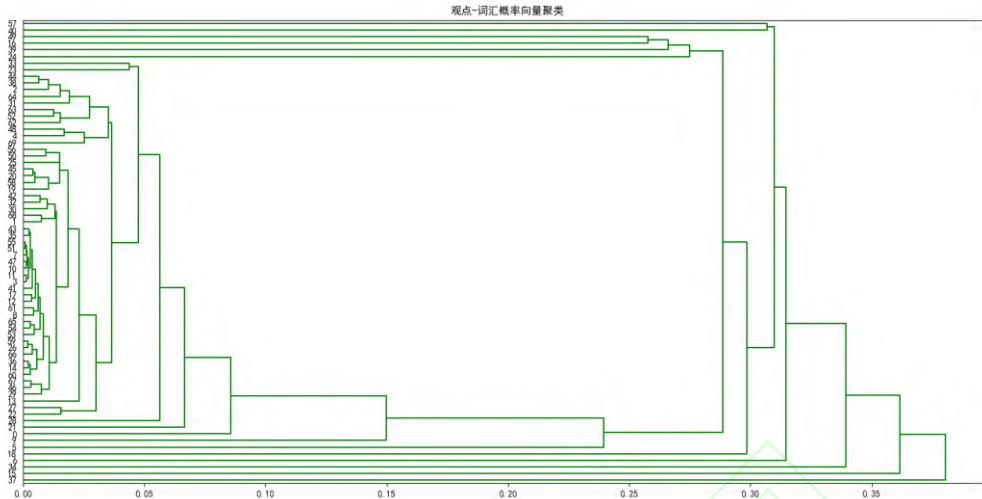


图10 观点向量聚类结果

Fig.10 Clustering results of view vectors

基于观点词汇概率向量和观点聚类结果，挑选出含义明确的聚类簇，构建了一个双层结构的城市规划舆情观点分类体系，其中包括23个观点小类和6个观点大类，并在图11中展示了其分类体系层次结构。

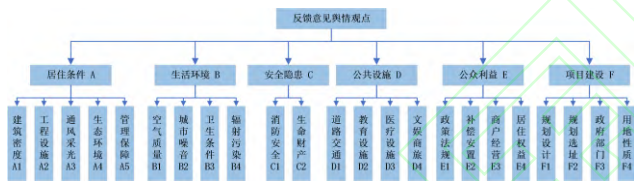


图11 城市规划网络舆情观点分类体系

Fig.11 The opinion classification system of network public opinion in urban planning

3.3 城市规划案例示例

为深入阐述本文提出的城市规划案例表达框架，本文以“住宅改建电梯间”城市规划案例为例，对舆情信息进行提取后按照城市规划网络舆情描述框架进行量化表达。

“住宅改建电梯间”城市规划案例的属性如表5所示。案例的时间属性向量为(2019,01,01,00,00,00)，表明规划项目公示的时间为2019年1月1日。空间属性中案例类别语义的向量维度为300，案例位置语义的向量维度为50。人口属性中年齡特征向量中维度值不小于0.08的年龄段有10-14岁、15-19岁、20-24岁、25-29岁、40-44岁。受教育程度特征向量的第四维度值最大，代表高中学历的人群所在比例最多。主要生活来源特征向量的第一维度值最大，表示近半数人的生活来源是劳动收入。

表5 城市规划案例属性表达示例

Tab.5 An example of attribute expression of urban planning case

名称		表达形式
基本信息	案例编号	1
	案例名称	住宅改建电梯间
时间属性	规划公示起始时间	(2019,01,01,00,00,00)
	案例类别	(0.224592,-0.264534,0.077634,0.348315, 0.0552,...)
空间属性	案例位置	(0.330404, 0.021668, -0.729113, 0.0127639, ...)
	年龄	(0.014,0.045,0.046,0.089,0.096,0.086,0.08 0,0.064,0.075,0.082,0.075,0.052,0.059,0.04 1,0.028,0.023,0.023,0.014,0.006,0.002,0.00 0,0.000)
人口属性	受教育程度	(0.013, 0.142, 0.319, 0.357, 0.104, 0.057, 0.005)
	主要生活来源	(0.460, 0.293, 0.001, 0.009, 0.002, 0.202, 0.028)

“住宅改建电梯间”城市规划案例的问题及结果如表6所示。根据3.2.2中的关注人群特征，可用28维向量表示该案例的关注人群。“住宅改建电梯间”城市规划案例的关注人群有业主、居民、孩子、学生、老人、领导、群众、病人、医生，可以看出对于电梯改建从幼至老都密切关注。该案例的情感倾向特征向量为(0.05, 0.04, 0.92)，呈现出强烈的负面情感倾向，表明约92%的居民对该规划项目持反对

意见。

本文定义案例的观点级情感倾向包括中性(0)、正面(1)、负面(2)、未提及(3)四种类型。案例的观点大类情感特征为(2,2,3,2,3,0)，可以看出居住条件、生活环境、公共设施方面的情感倾向以负面为主，项目建设方面的情感倾向以中性为主，而安全隐患和公众利益这两类观点在该案例中未提及到。

表6 城市规划案例问题及结果表达示例

Tab.6 An example of problem and result expression of urban planning case

案例问题	案例结果		表达形式
人群特征	粗粒度	案例关注人群	(1,1,1,1,1,1,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0)
	舆情特征	案例情感倾向	(0.05,0.04,0.92)
观点特征	细粒度	案例观点级	观点大类: (2,2,3,2,3,0)
	舆情特征	情感倾向	观点小类: (2,3,3,2,3,3,2,3,3,2,3,2,2,3,3,3,3,3,3,3,0,3,0,3)

4 结论

本文提出了一种城市规划案例描述模型和表达框架，实现了城市规划中地理信息和舆情信息的抽象化描述与量化表达。通过定义城市规划案例的时间属性、人口属性等地理环境特征和观点类别、情感倾向等网络舆情特征，有助于从现实地理环境和网络舆情观点等多个特征维度定量度量城市规划案例之间的相似度，为实现未来规划案例舆情发展态势的事前预测提供理论支持。结合关注人群和观点类别的舆情预测结果，可以有针对性的对未来规划项目建设方案进行提升，提高城市居民的满意度，从而为规划制定提供科学合理的决策支持。

未来的研究工作可围绕以下方面展开：(1) 观点类别的更新。在城市发展的不同阶段，城市居民的关注类型也在动态变化，需研究观点体系自动更新的方法。(2) 案例属性的扩充。城市规划案例舆情的影响因素复杂，未来的研究过程中可以在案例属性表达中融入更多的舆情影响因素，以科学准确地度量案例相似性。

参考文献

- [1] Mouratidis K. Urban planning and quality of life: A review of pathways linking the built environment to subjective well-being[J]. *Cities*, 2021, 115: 103229.
- [2] Li W, Feng T, Timmermans H J P, et al. Analysis of citizens' motivation and participation intention in urban planning[J]. *Cities*, 2020, 106: 102921.
- [3] Batty M, Axhausen K W, Giannotti F, et al. Smart cities of the future[J]. *The European Physical Journal Special Topics*, 2012, 214(1): 481-518.
- [4] Fung A. Putting the public back into governance: The challenges of citizen participation and its future[J]. *Public Administration Review*, 2015, 75(4): 513-522.
- [5] Koch P. Bringing power back in: collective and distributive forms of power in public participation[J]. *Urban Studies*, 2013, 50(14): 2976-2992.
- [6] Mahdavejad M, Amini M. Public participation for sustainable urban planning in case of Iran [J]. *Procedia engineering*, 2011, 21: 405-413.
- [7] Rall E, Hansen R, Pauleit S. The added value of public participation GIS (PPGIS) for urban green infrastructure planning[J]. *Urban Forestry & Urban Greening*, 2019, 40: 264-274.
- [8] Liu Chunju, Shi Dundun. The Research of Temporal Spatial Data Organization for the Urban Planning[J]. *Journal of Geo-information Science*, 2006, (04): 60-64. (刘春菊, 石敦敦. 城市规划信息系统的时空数据组织方案[J]. *地球信息科学*, 2006, (04): 60-64.)
- [9] Yang yansheng, Li Hailong. A Decision-Oriented Integration of Urban Planning Information[J]. *Urban Development Studies*, 2012, 19(11): 142-145. (杨言生, 李海龙. 面向决策支持的城市规划信息资源整合方法探讨[J]. *城市发展研究*, 2012, 19(11): 142-145.)
- [10] FREDERICO T, EGENHOFER M. Ontology-driven Geographic Information Systems [C] // MEDEIROS C B. Proceedings of the 7th ACM International Symposium on Advances in Geographic Information Systems. Kansas City, MO: ACM, 1999: 14-19.
- [11] Li Deren, Cui Wei. Geographic Ontology and SIMG[J]. *Acta Geodaetica et Cartographica Sinica*, 2006, 35(2): 143-148. (李德仁, 崔巍. 地理本体与空间信息多级网格[J]. *测绘学报*, 2006, 35(2): 143-148.)
- [12] Lü Guonian, Yuan Linwang, Yu Zhaoyuan. Surveying and Mapping Geographical Information from the Perspective of Geography[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10): 1549-1556. (闰国年, 袁林旺, 俞肇元. 地理学视角下测绘地理信息再透视[J]. *测绘学报*, 2017, 46(10): 1549-1556.)

- [13] Guo Renzhong, Lin Haojia, He Biao, et al. GIS Framework for Smart Cities[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(12): 1829-1835. (郭仁忠, 林浩嘉, 贺彪, 等. 面向智慧城市的 GIS 框架[J]. 武汉大学学报(信息科学版), 2020, 45(12): 1829-1835.)
- [14] Dong Weihua, Liu Yilong, Hei Qiaosong, et al. Research Framework of the Theories and Methods for Pan-Map Spatial Cognition[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(12): 2007-2014. (董卫华, 刘毅龙, 黑巧松, 等. 泛地图空间认知理论与方法研究框架[J]. 武汉大学学报(信息科学版), 2022, 47(12): 2007-2014.)
- [15] Wang Juanle, Zhang Min, Han Xuehua, et al. Spatio-temporal evolution and regional differences of the public opinion on the prevention and control of COVID-19 epidemic in China. *Acta Geographica Sinica*. 2020, 75(11): 2490-2504. (王卷乐, 张敏, 韩雪华, 等. COVID-19 疫情防控中的中国公众舆情时空演变特征. 地理学报, 2020, 75(11): 2490-2504.)
- [16] Dong Jianfeng. Study on Online Public Opinion Analysis Oriented Public Crisis Pre-warning [D]. Wuhan: Wuhan University, 2013. (董坚峰. 面向公共危机预警的网络舆情分析研究[D]. 武汉: 武汉大学, 2013.)
- [17] Zhou Zhipeng, Zhou Xingnan, Qian Lingfei. Online Public Opinion Analysis on Infrastructure Megaprojects: Toward an Analytical Framework[J]. *Journal of Management in Engineering*, 2021, 37(1): 04020105.
- [18] Du Yunyan, Zhou Chenghu, Shao Quanqin, et al. Theoretic and Application Research of Geo-Case Based Reasoning. *Acta Geographica Sinica*. 2002, 057(002): 151-158. (杜云艳, 周成虎, 邵全琴, 等. 地理案例推理及其应用. 地理学报, 2002, 057(002): 151-158.)
- [19] Yang Xinzong. Study of the Geographical Spatio-Temporal Process Representation Model Case-Based [D]. Shandong: Shandong University of Science and Technology, 2013. (杨新忠. 基于案例的地理时空过程表达模型研究[D]. 山东: 山东科技大学, 2013.)
- [20] Wen Wei. Spatial analysis methods and applications based on geographical cases [D]. Shandong: Shandong University of Science and Technology, 2009. (温伟. 基于地理案例的空间分析方法及应用[D]. 山东: 山东科技大学, 2009.)
- [21] Cheng Hu. Study on Public Participation and Social Management Mechanism[J]. *Administrative Law Review*, 2012, (01): 66-73+117. (程琥. 公众参与社会管理机制研究[J]. 行政法学研究, 2012, (01): 66-73+117.)
- [22] Hu Dingli, Li Rui, Meng Yao, et al. China's Urban Network from the Perspective of Toponym Co-occurrences in the News[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(02): 281-288. (胡定利, 李锐, 孟瑶, 等. 新闻地名共现视角下的中国城市网络[J]. 武汉大学学报(信息科学版), 2020, 45(02): 281-288.)
- [23] Jia Zheming, Li Wenjing. Governmental Digital Tools: Citizen Usage and Effects on Satisfaction of Government. *Journal of Beijing University of Aeronautics and Astronautics(Social Sciences Edition)*, 2017, (02): 4-12. (贾哲敏, 李文静. 政务新媒体的公众使用及对政府满意度的影响. 北京航空航天大学学报(社会科学版), 2017, (02): 4-12.)
- [24] Du Yunyan, Wen Wei, Cao Feng. A study on spatial data mining using Geo-CBR and its application. *Geographical Research*, 2009, 28(05): 1285-1296. (杜云艳, 温伟, 曹锋. 空间数据挖掘的地理案例推理方法及试验. 地理研究, 2009, 28(05): 1285-1296.)
- [25] Yao Di, Zhang Chao, Huang Jianhui, et al. Semantic Understanding of Spatio-Temporal Data: Technology & Application. *Journal of Software*, 2018, 029(007): 2018-2045. (姚迪, 张超, 黄建辉, 等. 时空数据语义理解: 技术与应用. 软件学报, 2018, 029(007): 2018-2045.)
- [26] Guo Shuiqiong. Study on the convergence of land use classification between urban planning and land use planning[C]// Guiyang, Guizhou. China Urban Planning Annual Conference, 2015:176-188. (郭谁琼. 城市规划和土地利用规划用地分类衔接研究[C]//新常态: 传承与变革——2015 中国城市规划年会论文集 (09 城市总体规划). 贵州贵阳. 中国城市规划年会, 2015: 176-188.)
- [27] Chen Fu, Chen Haiyan, Zhu Zhenhua, et al. Analysis on Evaluation of Urban Residential Quality and Satisfaction[J]. *Human Geography*, 2000, (04): 24-27+13. (陈浮, 陈海燕, 朱振华, 等. 城市人居环境与满意度评价研究[J]. 人文地理, 2000, (04): 24-27+13.)
- [28] Thet T T, Na J C, Khoo C S G. Aspect-based sentiment analysis of movie reviews on discussion boards[J]. *Journal of Information Science*, 2010, 36(6): 823-848.

网络首发:

标题: 基于地理案例的城市规划信息组织

作者: 李锐, 王璟琦, 王顺利, 李江, 吴华意

DOI: 10.13203/j.whugis20220395

收稿日期：2023-05-04

引用格式：

李锐，王璟琦，王顺利，等. 基于地理案例的城市规划信息组织[J]. 武汉大学学报（信息科学版），2023，DOI: 10.13203/j.whugis20220395（LI Rui, WANG Jingqi, WANG Shunli, et al. Information Organization in Urban Planning Based on Geographical Case[J]. *Geomatics and Information Science of Wuhan University*, 2023, DOI: 10.13203/j. whugis20220395）

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：

基于地理案例的城市规划信息组织

朱杰，游雄，夏青，张宏军

武汉大学学报（信息科学版），2020, 45(7): 1027-1034

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20180313>

