

基于本体的产业园区价值评估模型研究

吴瑞龙¹, 何华贵¹, 陈朝霞¹, 李少智¹, 周 勍¹, 刘 洋^{1,2}

(1. 广州市城市规划勘测设计研究院, 广州 510060; 2. 广东省城市感知与监测预警企业重点实验室, 广州 510060)

摘要:现有产业园区价值评估停留在人工评分模式,大多基于传统的数理统计,不够智能化和规范化。针对现有技术不足,利用本体论相关技术,首先定义产业园区的本体概念,围绕产业园区价值评估的影响因素,抽象并定义产业园区相关类的概念、对象关系以及属性,建立产业园区的领域本体模型,形成产业园区知识本体,在此基础上构建产业园区价值评估的计算规则,建立产业园区的价值评估模型。研究成果可针对不同类型的产业园区进行量化价值评估,为产业园区价值评估研究提供思路,具有一定参考意义。

关键词:产业园区;价值评估;领域本体;知识模型

中图分类号:F279.24 **文献标志码:**A **文章编号:**1671-1807(2022)05-0159-05

产业是城市经济体系的基础和核心,经济与产业发展是支撑一个城市可持续发展的基础^[1-2]。产业园区是指为促进产业发展为目标而创立的特殊区位环境,是区域经济发展、产业调整升级的重要空间聚焦形式^[3],承载各类市场主体开展产业活动和创新活动的重要载体,是决定城市发展质量的基本单元,担负着聚集创新资源、培育新兴产业、推动城市化建设等一系列的重要使命^[4-5]。对产业园区的价值进行评估尤为重要,可以取长补短,对于劣势的因素进行补强,可作为产业园区调整升级发展的重要决策依据,也是城市创新经济发展的必然要求^[6-7]。

现有关于产业园区评估主要通过专家评分模式,没有形成规范的评估体系,一般都是针对某种特定类型的产业园区进行评估,如文献[8]中通过选取经济发展质量、创新创业活力、对外开放水平、资源配置效率、管理运营质量、环境友好质量等 6 个维度,43 个代表性指标进行评估。这类评估对于信息采集依赖度较高,当信息完整度不高或者信息缺失时,较难支撑价值评估,同时评估方法基于传统数理统计模型,不够智能化,而且只能针对某一类产业园区进行评估,对于其他类型产业园区可复用性不强。本文将本体论相关知识应用于产业园区的价值评估,通过对产业园区以及影响其价值评估

的因素进行抽象和定义,对产业园区进行本体建模,形成产业园区的知识本体,并通过构建的价值评估规则对产业园区进行价值评估,具有可复用性、方便扩展的特点。

1 本体概念及表示语言

本体(ontology)的概念包含两层含义:一层是从哲学层面,其源自哲学学科,是哲学的一个分支,主要是研究存在的本质,即对客观存在的系统或事物的解释说明,是一种描述术语及术语间关系的概念模型,本体关注的是客观现象的抽象本质;另一层意思是指将其引申应用到计算机科学领域和信息科学领域内,可作为一种“形式化的,对于共享概念体系明确而又详细的说明”,其核心内容是指一种描述由一套对象类型(概念或类)、属性以及关系所构成的整体模型^[9-10]。随着人工智能的发展,在人工智能领域被赋予了新的定义,从语义网的角度来理解本体的概念,可被定义为组成领域的基本术语及其之间的关系,本体论作为语义网的重要组成部分,能够在语义信息和知识层次对系统进行描述并进行建模^[11-13]。

为了使本体被计算机理解,通常采用形式化描述语言来表示本体,这种语言具有一定的推理能力,在互联网共享信息高速发展的背景下,根据本体在应用中的不同作用,本体所使用的语言也不

收稿日期:2022-01-14

基金项目:广东省城市感知与监测预警企业重点实验室基金项目(2020B121202019);广州市工信委项目(GZIT2016-A5-147);广州市城市规划勘测设计研究院基金项目(RDI2210201065)。

作者简介:吴瑞龙(1990—),男,河南郑州人,广州市城市规划勘测设计研究院,工程师,硕士,研究方向为知识图谱、空间信息智能服务研发、分布式地理计算。

同,本体描述语言主要是为了将本体的逻辑概念模型以文件编码的方式展现出来,使用户将领域模型进行清晰的、形式化的概念描述要求,因此对于本体描述语言需要满足以下要求:定义良好的语法和语义、有效的推理支持、充分的表述能力、表达的便利性。OWL(Web Ontology Language)是由 W3C 提出的标准的本体描述语言,是一种知识表示语言,对领域知识进行明确表达推理,其目的是捕获知识,有以下 3 个基本概念^[14-15]:

- 1)公理(axioms):一个 OWL 本体表达的基本陈述。
- 2)实体(entities):用来指向现实世界对象的元素。
- 3)表达式(expressions):实体的组合,从简单的描述形成复杂的描述。

2 产业园区领域本体构建

2.1 产业园区语义描述

在对产业园区进行描述时,一般都会包含其建设时间、假设背景、发展特色、区位优势、产业优势、政策资源等众多共性因素。比如百度百科对于深圳高新技术产业开发区的描述:深圳高新技术产业开发区始建于 1996 年 9 月,深圳高新技术产业开发区位于深圳市南山区位于广东省深圳经济特区西

部,行政区域东起车公庙与福田区相邻,2008 年,高新区在占全市不到 0.6%的土地上,实现工业总产值 2 249.78 亿元,同比增长 17.60%,占全市工业总产值的 14.19%,高新区每平方公里工业总产值 196 亿元、高新技术产品产值 186 亿元、工业增加值 43 亿元。再比如广州经济技术开发区的描述:广州经济技术开发区于 1984 年经国务院批准成立,是全国首批国家级经济技术开发区之一,地处广州市东部。与广州高新技术产业开发区广州出口加工区、广州保税区、中新广州知识城合署办公(统称“广州开发区”),实行“五区合一”的管理体制。为加速广州高新技术产业的发展,1997 年广州市政府对高新区管理体制进行了调整,形成由广州科学城、天河科技园、黄花岗科技园和民营科技园组成的“一区多园”的新格局。开发区交通便利,生态环境良好,基础设施一应俱全,投资环境优越。

2.2 产业园区本体类定义

通过对产业园区的语义描述,以及其价值影响因素的语义描述分析,围绕产业园区的价值评估的目的,抽象并定义了产业园区相关的本体类和影响其价值评估因素的类,分别是区位类、活力类、创新类、成本类,对每个类的定义、对象关系、类关系、属性给出详细说明,见表 1。

表 1 产业园区本体类定义

类	定义	实例	对象关系	类关系	对象属性
产业园区	继承自 owl:Thing,表示产业园区本体,描述产业园区的术语构成及其关系	广州经济技术开发区、深圳高新技术产业开发区	has_区位、has_活力、has_创新、has_成本	owl:Thing 的子类	园区位置、园区等级、占地面积、建造时间、园区类型

表 1 中 owl:Thing 表示所有事物的父类,对象关系“has_区位”表示对象“产业园区”和对象“区位”是“has”(拥有)的关系,其他类比。

2.2.1 区位类

区位是对产业园区的地理位置的优势进行描述的表达,对于产业园区的价值评估具有普遍性

意义,一般用距离值来衡量,通常在描述产业园区区位的时候主要包括 3 个方面,包括交通便利性、创新源邻近度、到核心城市的距离,对此本文抽象出 3 个子类,分别是到大型交通枢纽距离、到核心城市距离、到创新机构距离,其类的详细描述构成见表 2。

表 2 区位类定义

类	定义	实例	类关系	属性
区位	描述区位的术语构成		owl:Thing 的子类	因子系数
到大型交通枢纽距离	描述交通便利性的类	到火车站距离、到机场距离	区位的子类	
到核心城市距离	描述到核心城市距离的类	到某一线城市的距离为、到二线城市的距离	区位的子类	
到创新机构距离	描述到创新机构距离的类	到最近高校的距离、到最近科研机构的距离	区位的子类	

2.2.2 活力类

活力类是对产业园区的产业活跃度进行描述的表达,一定程度体现了产业园区发展的能力和潜

力,对于产业园区的价值评估具有重要参考意义,可用园区内各个企业的行业类别、企业规模以及园区内就业人员等因素进行评估,为此抽象出 3 个类,

包括企业行业、企业规模、就业人员,详细描述构成 见表 3。

表 3 活力类定义

类	定义	实例	类关系	属性
活力	描述活力的术语构成		owl:Thing 的子类	因子系数
企业行业	描述园区内企业的行业的类	金融业、交通运输业	活力的子类	
企业规模	描述产业园区内企业注册资本规模的类	100 万~500 万、500 万~1 000 万	活力的子类	
就业人员	描述产业园区内就业人员的类	青年就业人员占比、通勤人员数量	活力的子类	

2.2.3 创新类

产业园区的管理模式和运营模式很大程度决定了产业园区的长远发展,尤其体现在创新管理和创新运营,对产业园区的未来发展有较大的影响,

是产业园区发展和价值体现的重要因素,主要包括两个方面:高新技术企业和科技成果产出,相关类的详细描述构成见表 4。

表 4 创新类定义

类	定义	实例	类关系	属性
创新	描述创新的术语构成		owl:Thing 的子类	因子系数
高新技术企业	描述产业园区高新技术企业的类	电子信息科技企业、航空航天企业	创新的子类	
科技成果	描述产业园区内科技成果产出的类	到最近高校的距离、到最近科研机构的距离	创新的子类	

2.2.4 成本类

成本是产业园区发展的重要制约因素,是提高产业园区建设经济效益的重要手段,如何更好地控制产业园区的运营成本,获取更大的经济效益,是

众多产业园区面临的主要问题,将其作为一项重要的价值评估因素,可为产业园区的成本控制提供参考,主要包括 4 个方面:人工成本、租金成本、通勤成本、生活成本,其类的详细描述构成见表 5。

表 5 成本类定义

类	定义	实例	类关系	属性
成本	描述成本的术语构成		owl:Thing 的子类	因子系数
人工成本	描述产业园区人工成本的类	员工工资	成本的子类	
租金成本	描述产业园区租金成本的类	写字楼租金、园区厂房租金	成本的子类	
通勤成本	描述产业园区内通勤成本的类	产业园区通勤距离,通勤交通便利性评分	成本的子类	
生活成本	描述在产业园区内生活的成本	水电气价格成本评分	成本的子类	

通过以上抽象类及其相关概念的定义,形成产业园区及其相关价值评估因素的类关系图,如图 1 所示。

2.3 产业园区本体构建

根据产业园区及其相关价值评估类的定义,使用 Protege 软件构建产业园区的本体知识模型。Protege 是基于 Java 语言开发的开元的本体编辑和知识获取软件,是一款本体开发工具,也是基于知识的编辑器,在本体类的基础上形成产业园区本体模型的 OWL 文档描述,形成的本体知识类图如图 2 所示。

3 产业园区价值评估计算

产业园区价值评估计算主要利用产生式规则,其结构包括前提和结论两部分:前提(或 IF 部分)描述状态和结论(或 THEN 部分)描述在前提状态存在的条件下所执行的操作,产生式规则一般指形如

$\alpha \rightarrow \beta$ 或 IF α THEN β 或其等价形式的一条规则,其中 α 代表产生式的前件,一般由逻辑组合或者逻辑表达式来构成, β 表示产生式的后件,表示若干个结论或动作,表示在满足 α 先决条件情况下,推出的结论或应执行的操作,可以将产生式推理描述为“当前提条件 α 满足时,则可以推出结论 β' ”。

在本文中,将正则表达式抽取的产业园区语料短语作为先决条件 α , α 由上一步骤中对产业园区本体 OWL 文档的解析得到,并作为正则表达式的匹配模板,其中语料短语是已经包含了对于某个因素评分的短语,该评分通常由外部专家综合评分打分,如“深圳高新技术产业开发区始建于 1996 年 9 月,位于深圳市南山区位于广东省深圳经济特区西部,对外交通便利综合评分 85”,产业园区价值评估计算就是判断在条件 α 时,应该采取哪种计算规则 β ,构造的价值评估计算规则见表 6。

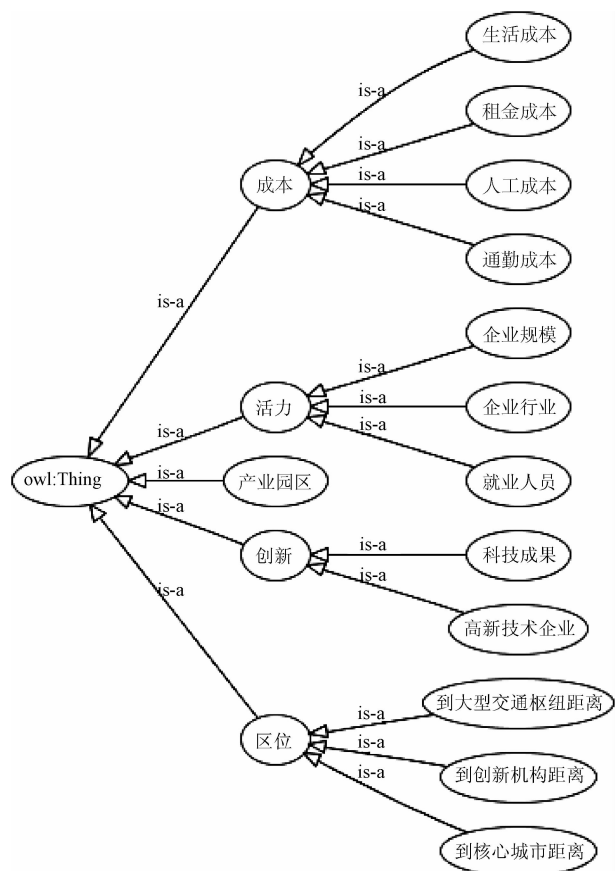


图 1 产业园区类关系图

表 6 中 Num 代表组成该项评分的量值,由阿拉伯数字组成,Num:(0|1|2|3|4|5|6|7|8|9|),V 代表对产业园区的价值评估得分。对于输入的产业园区语料信息,利用从产业园区本体 OWL 解析得到的正则表达式模板提取文本中相关的语料短

语,然后查询数据库所对应的计算规则,构建 SWRL 描述规则(SWRL-Semantic Web Rule Language 是由以语义的方式呈现规则的一种语言),然后将规则读入 Jena 推理机进行价值评估计算,将价值评估结果返回。对抽取的产业园区语料短语构造如下 SWRL 计算规则:

$\text{ChaneyuanquPhrase} (? x) \wedge \text{has_qw_01} (? x, ? qw1) \wedge \text{has_hl_01} (? x, ? hl1) \wedge \text{has_cx_01} (? x, ? cx1) \rightarrow \text{calculateValue} (? x, \text{RuleThen})$

其中,ChaneyuanquPhrase(? x)表示 x 是输入的产业园区语料短语,has_qw_01(? x,? qw1)表示 x 具有区位编号 QW_01 的得分为 qw1,符号“^”表示后续为并列条件,calculateValue(? x,RuleThen)表示可以根据 RuleThen 计算结果,在对时间短语 x 进行计算时选择 RuleThen 类型的计算规则。举例如下,假设有输入语料:“深圳高新技术产业开发区始建于 1996 年 9 月,位于深圳市南山区位于广东省深圳经济特区西部,对外交通便利综合评分 85,综合得到企业规模评分 88,2020 年通过评审得到科技成果评分 80”。根据规则库,可以抽取出 3 个产业园区价值评估规则分别是:“对外交通便利综合评分 85”“企业规模评分 88”“科技成果评分 80”,然后根据计算规则 RuleThen 进行计算, $V = 85 \times 0.4 + 88 \times 0.3 + 80 \times 0.5 = 100.4$

4 结论与展望

本文采用本体论的相关技术,对产业园区进行本体建模,定义了产业园区本体类的相关概念和定义。围绕产业园区价值评估的一般描述进行抽象,

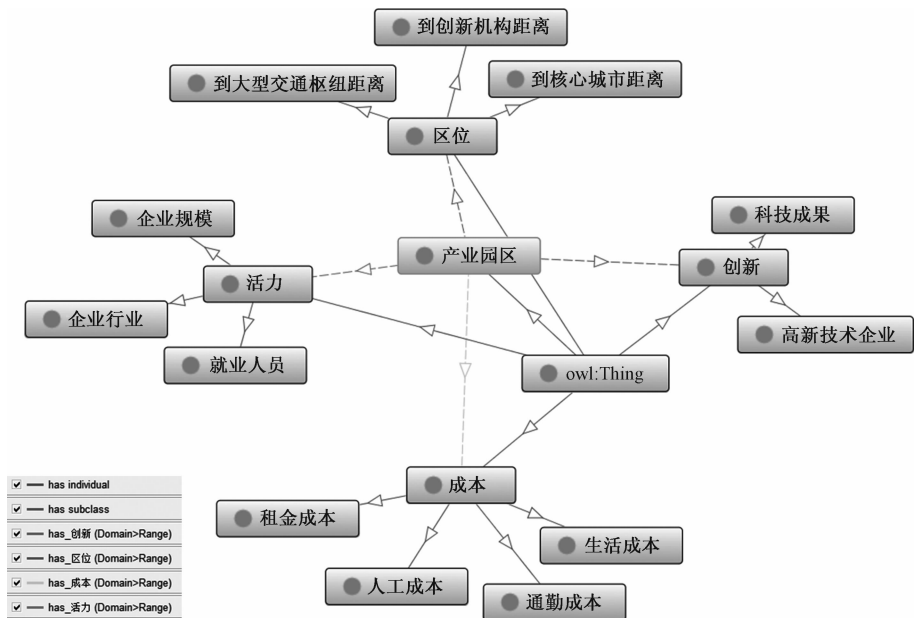


图 2 产业园区类构成图

表 6 产业园区价值评估计算规则

规则编号	规则类别	规则名称	RuleIF	因子系数	RuleThen
QW_01	区位	has_qw_01	对外交通便利性综合评(Num{1,2})	0.4	V=V+因子系数× (Num{1,2})
QW_02	区位	has_qw_02	到核心城市的距离评分(Num{1,2})	0.3	
QW_03	区位	has_qw_03	到创新机构距离评分(Num{1,2})	0.3	
HL_01	活力	has_hl_01	企业行业评分(Num{1,2})	0.3	
HL_02	活力	has_hl_02	企业规模评分(Num{1,2})	0.3	
HL_02	活力	has_hl_03	就业人员评分(Num{1,2})	0.4	
CX_01	创新	has_cx_01	高新技术企业评分(Num{1,2})	0.5	
CX_02	创新	has_cx_02	科技成果评分(Num{1,2})	0.5	
CB_01	成本	has_cb_01	人工成本(Num{1,2})	0.25	
CB_02	成本	has_cb_02	租金成本(Num{1,2})	0.25	
CB_03	成本	has_cb_03	通勤成本(Num{1,2})	0.25	
CB_04	成本	has_cb_04	生活成本(Num{1,2})	0.25	

分别定义了区位类、活力类、创新类、成本类,形成了产业园区价值评估的知识模型,形成了产业园区的价值评估的可复用知识模型,且模型具备可扩展性。依据定义的产业园区本体模型,设计了产业园区价值评估的计算规则,可以根据输入的语料信息利用正则表达式提取相关预料信息,并按照计算规则对产业园区进行价值评估,而不是传统机械的评估方式对产业园区进行价值评估,减少了数据的依赖性,形成了较为规范的产业园区价值评估体系,可以对不同类别的产业园区进行价值评估。但本文研究也存在不足,如没有建立较为全面的产业园区知识本体库,在后续研究中将结合图数据库等技术建立产业园区的知识图谱,为产业园区价值评估提供更加全面的支撑。

参考文献

[1] 乔朝飞,王硕.我国地理信息产业园发展现状分析及对策建议[J].测绘与空间地理信息,2017,40(10):11-14.

[2] 方创琳,黄金川.高科技园区与城市协调发展的生态调控[J].地理学报,2003(3):390-397.

[3] 王玲杰.工业园区国土数据管理及分析系统的设计与实现[J].测绘与空间地理信息,2021,44(2):148-150,153,156.

[4] 张晓平.我国经济技术开发区的发展特征及动力机制[J].

地理研究,2002(5):656-666.

[5] 俞孔坚,黄国平.高科技园区的发展战略与生态规划:以广东省南海信息产业园为案例[J].城市发展研究,2001(3):58-64.

[6] 陈忠暖,陈汉欣,冯越,等.新世纪以来广东文化产业发展与演变:与国内文化大省的比较[J].经济地理,2012,32(1):76-84.

[7] 刘轩朋,朱富晓.基于 ICT 技术的智慧化工园区数字化管理模式研究:以衢州绿色产业集聚区为例[J].测绘与空间地理信息,2021,44(6):132-134.

[8] 刘佳骏.中国海外合作产业园区高质量建设评价体系研究[J].国际经济合作,2021(3):59-67.

[9] 黄映辉,李冠宇.Ontology 的实质是“本体论模型”[J].计算机工程与应用,2007,43(23):12-14.

[10] 汤新明.形式概念分析理论研究及在面向对象程序中的应用[D].鞍山:辽宁科技大学,2008.

[11] 郑丽萍.本体映射的研究[D].青岛:山东科技大学,2005.

[12] 鲍晓.领域本体概念及概念间关系学习算法研究[D].武汉:华中科技大学,2013.

[13] 徐红升.基于形式概念分析的本体构建、合并与展现[D].开封:河南大学,2007.

[14] 郭嘉琦.领域本体的构建及其在信息检索中的应用研究[D].北京:北京邮电大学,2007.

[15] 沈国华.基于描述逻辑的语义 Web 服务建模及推理研究[D].南京:南京航空航天大学,2009.

Research on Value Evaluation Model of Industrial Park Based on Ontology

WU Ruilong¹, HE Huagui¹, CHEN Zhaoxia¹, LI Shaozhi¹, ZHOU Qing¹, LIU Yang^{1,2}

(1. Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510060, China;
2. Guangdong Enterprise Key Laboratory for Urban Sensing, Monitoring and Early Warning, Guangzhou 510060, China)

Abstract: Existing Industrial Park value evaluation stays in the manual scoring mode, which is mostly based on traditional mathematical statistics, and is not intelligent and standardized. In view of the shortcomings of existing technologies, using ontology related technologies, the ontology concept of Industrial Park is defined, the concepts of related categories of Industrial Park is abstracted and defined around the influencing factors of Industrial Park value evaluation based on the object relationship and attribute, the domain ontology model of Industrial Park is established to form the knowledge ontology of Industrial Park. On this basis, the calculation rules of Industrial Park value evaluation are constructed, and the value evaluation model of Industrial Park is established. The research can carry out quantitative value evaluation for different types of Industrial Parks, which provides ideas for the research of Industrial Park value evaluation, it has certain reference significance.

Keywords: Industrial Park; value assessment; domain ontology; knowledge model