

基于扩展语义匹配的 BIM 和 GIS 三维建筑数据融合

马 骏 胡瑛婷 石 玉 闫 伟

(山东师范大学 信息科学与工程学院, 济南 250300)

【摘 要】目前,在 BIM 与 GIS 融合方面的研究主要分为三个方向,分别是数据格式转换、数据标准扩展和本体论,本项目选择通过本体论的模式进行 IFC 和 CityGML 的实体融合。通过语义修正提取实体的三元组,再用知识表示学习将三元组表示为实体向量,将实体语义信息转换到低维向量空间,采取以扩展语义匹配为基础构建 IFC 和 CityGML 标准的匹配融合系统,解决在 IFC 和 CityGML 两个不同的标准下对同一实体的语义方面的理解偏差问题,并通过相似度计算进行聚类。此外还通过提取到的三元组,利用实体—关系—实体进行几何重构,将 IFC 标准的实体映射到 CityGML 标准的实体。最后用建立的三维模型验证,实现了实体从 BIM 领域到 GIS 领域的转换。

【关键词】BIM 技术; GIS 技术; IFC 标准; CityGML 标准; 知识表示学习; 几何重构; 融合

【中图分类号】TU17 **【文献标识码】**A

【版权声明】文集数据被中国知网重要会议论文全文数据库(CPCD)收录,被本刊录用并在中国知网网络首发正式出版,严禁侵权转载。

引言

伴随着信息技术的高速发展,应运而生的出现了如何将新一代信息技术创新性应用到优化城市结构等关键问题,而提升城市竞争力的有效方法之一就是智慧城市的建设。自 2009 年,中国连续召开了二十多场智慧城市研讨会之后,“智慧城市”这一理念开始受到群众和学者的广泛关注。从数字城市到智慧城市的进一步升级的一个特征是更加精细化的城市管理。为成功实现城市的进一步精细,仅利用建筑信息模型(building information modeling, BIM)会缺少在地理坐标上的数据信息,而如果仅利用地理信息系统(geographic information system, GIS)则会无法得到建筑物室内的的构件组成,因此 BIM 和 GIS 的融合则可结合二者互补的数据信息,将 BIM 领域中单栋建筑的精确模型数据信息和 GIS 领域中空间信息系统数据信息相结合,它不仅拓宽了

数据的表示范围,还增强了数据的深度,最终实现将从建筑物外走进建筑物内的目标。

在 BIM 技术的高速发展的潮流趋势下,三维城市模型提供了许多详细的数据信息,提高了城市管理效率^[12],同时也为模型的建立提供了更加丰富的数据支持^[13],是目前城市三维建模研究的潮流趋势^[2],也是智慧城市等相关研究领域的技术需求^[14]。

本文选取 IFC (Industry Foundation Classes) 和 CityGML (City Geography Markup Language) 两个标准为载体来研究 BIM 和 GIS 的融合。

IFC 标准采用 EXPRESS 语言定义建筑数据,反映了现实世界中的实体。CityGML 标准则是定义了城市中的地理位置数据信息等,并基于 XML 格式来完成数据的交换与存储。

本文提出了一种基于扩展语义匹配的 BIM 和 GIS 三维建筑数据融合,采取了知识表示学习方法,

【基金项目】 国家自然科学基金资助项目(编号:62002207; 62072290; 62073201); 山东省自然科学基金资助项目(编号:ZR2020MA102)

【作者简介】 马骏(2000-),男,本科生在读,主要研究方向:BIM 和 GIS 匹配融合;胡瑛婷(2000-),女,本科生在读,主要研究方向:BIM 和 GIS 匹配融合;石玉(2000-),女,本科生在读,主要研究方向:BIM 和 GIS 匹配融合;通讯作者:闫伟,女,副教授、硕士生导师,主要研究方向:知识建模、知识融合、知识计算、智慧城市等。

先提取三元组,将三元组拆分成实体和关系两部分,将实体部分转换到低维向量空间,将语义信息表示为实体向量,再对向量进行聚类。此外还采用几何重构的方法,将知识表示学习无法实现映射的实体进一步的完成映射对应。最后建立三维模型,在实例层面进行验证。

1 BIM 和 GIS 的融合方法研究现状

1.1 在学术界的 research 现状

BIM 与 GIS 的融合研究是当前相关专业研究的前沿和趋势,针对此方面,许多学者和组织都进行了一定的探讨和研究^[1]。

Shinji Yamamura 等人^[3]提出了一种基于 GIS-BIM 的城市能源布局体系,提供了城市能源规划技术和处理方案。B. Wu 等人^[4]分析了 GIS 和 BIM 两种技术融合带来的好处,讨论并研究了优化 GIS 和 BIM 数据集之间转换的方法。Yichuan Deng 等人^[5]通过实例生成了 IFC 和 CityGML 之间的映射规则。E. P. Karan 等人^[6]利用 Web 技术,成功的获取不同领域的的数据信息,解决 BIM 和 GIS 之间的数据共享问题。

而现在之所以没有实现 IFC 标准和 CityGML 标准的完整映射和数据的补充,主要是因为存在以下不足:

(1) 语义信息不对等

IFC 标准中包含的语义信息和数据信息比 CityGML 标准中的信息更多更详细,假如根据 CityGML 标准建立映射关系的话,那将必然损失大量数据,而解决数据丢失这一问题则还需要更进一步的研究。

(2) IFC 标准和 CityGML 标准对语义对象定义的侧重点不同。IFC 标准中更关注建筑物的构件。而 CityGML 标准则更侧重于空间属性。

(3) 不具备完整的映射规则

没有将两个标准数据本身作为出发点,同时考虑几何和语义信息。

1.2 在工业界的 research 现状

目前在商业软件方面,已经有成功实现了将 IFC 模型与多层 LOD 级别的 CityGML 模型的双向转换的软件,包括 FME、IFCExplorer, BIMServer, 等。其中,FME 可将 IFC 标准的模型转换为 CityGML 标准的模型。IFCExplorer 目前已实现导入 IFC 标准的

模型和 CityGML 标准的模型、并进行场景整合和模型规范验证等功能,但转换仍停留在从 IFC 标准的模型到低层次的 CityGML 标准的模型的转换研究上; BIMServer 可实现对 IFC 标准的模型管理,可实现导出 CityGML、KML 等格式,但存在问题是导出的 CityGML 标准的模型精度低。

为了实现的 BIM 和 GIS 两大技术领域的数据上的真正融合,在其载体 IFC 标准与 CityGML 标准数据格式转换主要分为以下几类:

(1) U. Isikdag^[7]提出了将 IFC 标准的模型转换为 CityGML 标准的模型的框架,该框架描述了实现从 IFC 标准到 CityGML 标准的 LOD1 - LOD4 映射的基本思路,但没有提出设计实现具体的映射方法,但只是思路而没有实践,且该规则是单方向的将 IFC 标准的模型映射为 CityGML 标准的模型,而无法实现反向的转换。

(2) 设计中间模型。M. El - Mekawy^[8]提出将 IFC 标准与 CityGML 标准的数据信息整合到一个中间模型中,将两个标准的实体整合到中间模型中。

1.3 本文研究方法

本文对于基于扩展语义匹配的 BIM 和 GIS 三维建筑数据融合,以 IFC 和 CityGML 标准为载体,研究路线分为两个方面,第一方面是建立 IFC 和 CityGML 的本体匹配模型,采用字符串匹配、知识表示学习、几何重构三个方法实现从 IFC 标准的实体到 CityGML 标准的实体的映射匹配。第二方面是实例验证,在三维模型层面验证本体匹配模型。

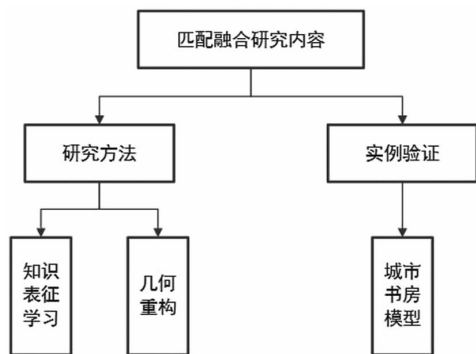


图1 研究内容

2 建立 IFC 和 CityGML 的本体匹配模型

2.1 知识表征学习

2.1.1 语义修正

提取实体的三元组,即对两个标准的文件进行

信息抽取。方法是通过 python 的 rdflib 库对 owl 文件提取三元组 (head、relation、tail)。我们在获得 IFC4.0 以及 CityGML2.0 的 owl 文件之后,通过 python 提供的现有库 rdflib 对这两个 owl 文件进行解析,抽取 IFC 以及 CityGML 的三元组,并且对异常数据进行处理,最终分别得到 IFC 和 CityGML 的三元组,并且为了方便后续的向量化处理,我们对实体,关系以及三元组进行人工标注处理。我们最终可以通过人工标注的实体,关系,三元组,共获得 2 527 个 CityGML 三元组,36 114 个 IFC 三元组。我们后来对这些提取的三元组进行异常数据的处理,获得 290 个 CityGML 实体以及 4 923 个 IFC 实体。

表 1 部分三元组

实体 1	关系	实体 2
IfcLengthMeasure_List	hasContents	IfcLengthMeasure
IfcFaceOuterBound	type	IfcFaceOuterBound
IfcWallStandardCase	name_IfcRoot	IfcLabel
_Opening	subClassOf	_CityObject
Window	type	Class
OuterFloorSurface	subClassOf	_BoundarySurface
textureCoordinates	type	DatatypeProperty
lod2Solid	type	ObjectProperty
CeilingSurface	type	Class
Door	subClassOf	_Opening

2.1.2 知识表示学习

(1) 利用 TransE 实现三元组的向量化处理,并对提取出的三元组进行知识表示。

采用了 Trans 家族的 TransE 算法^[9],实现将实体和关系用向量进行表示。利用实体向量,再进一步研究 IFC 和 CityGML 本体匹配融合。利用 TransE 方法,对提取的三元组进行向量化处理,将人工标注好的实体以及关系文件输入 TransE 模型,可以得到对应的实体以及关系的 50 维的向量。

在 TransE 算法中定义了如下超参:向量维度为 50,学习率为 0.01,目标函数常数为 1,使用 L1 范数计算距离。

实现将知识图谱中所有的实体、关系用向量来进行表示。原理如图 2 所示。

(2) 使用 T-SNE^[11] (t-distributed Stochastic Neighbor Embedding) 对该高维(50 维)向量降维并可视化处理。

①将上步得到的高维向量数据随机投影到一

维数轴中。

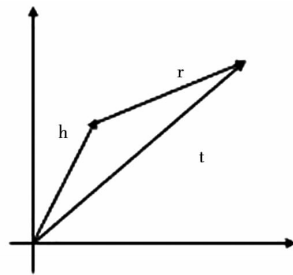


图 2 TransE 原理图

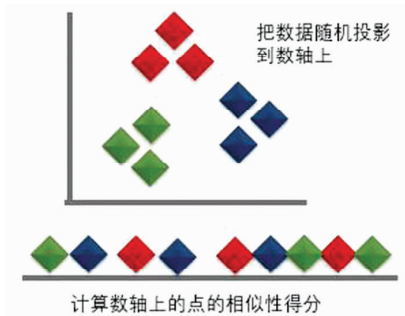


图 3 T-SNE 原理图(a)

②再使用 t 分布计算样本间的相似性并得出一个分数。

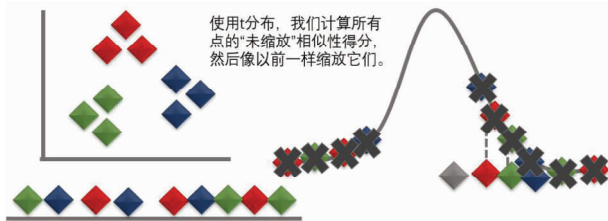


图 4 T-SNE 原理图(b)

③变换目标,并计算剩余样本和目标之间的相似性得分,最终得到一个相似性得分矩阵。

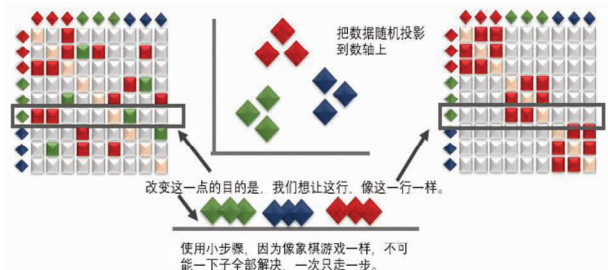


图 5 T-SNE 原理图(c)

④T-SNE 的目的是将样本逐步移动,最终完成原始聚类数据的恢复和还原。

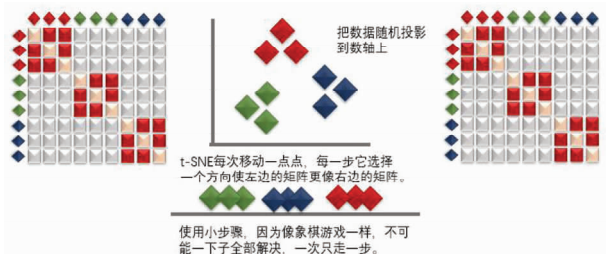


图 6 T-SNE 原理图(d)

2.1.3 相似度计算

得到实体向量后,为进一步对 IFC 和 CityGML 本体融合匹配,利用 k-means 聚类对实体向量进行进一步的分析。

k-means 聚类算法流程:

- (1)确定 IFC 和 CityGML 全部实体聚类结果 的类数 k 值。
- (2)随机抽取 k 个数据点作为质心。
- (3)遍历数据集中每一个实体点,计算其与每一个质心的间隔距离,判断其与哪个质心距离最近,就划分到哪个质心所属于的集合。
- (4)把全部数据归好类后,会获得 k 个集合。再重新计算每个集合的质心。
- (5)假如新得到的质心和原本的质心之间的间隔距离小于阈值,则认为聚类结束。
- (6)若新质心和原质心的间隔距离变化大,则还需不断迭代③~⑤步骤。

通过对 IFC 和 CityGML 实体向量进行 K-means 聚类,聚类结果如表格 2 和图 6 所示,我们将聚类中心用绿点表示。

表 2 IFC 和 CityGML 实体聚类结果

IFC 实体	CityGML 实体
ifcwindow, ifcdoor	door, window
ifcmateriallayerwithoffsets	transformationmatrix2x2type
opening	opening
innerboundaries_ifccurveboundedplane	interiorroom, consistsofbridgepart
ifcflightdistributiondata, elevation	_object, elevation
basesurface	interiorbridgeroom, basesurface
objectproperty, functionalproperty	objectproperty
material	simpletexture, material
ifcpresentationstyleassignment	xmlschemainteger, doublebetween0and1
polygon	polygon
ifcfillstyleselect	xmlschemanonnegativeinteger
height	height

2.2 几何重构

虽然门和窗户等都可通过用 2.2 知识表示学习

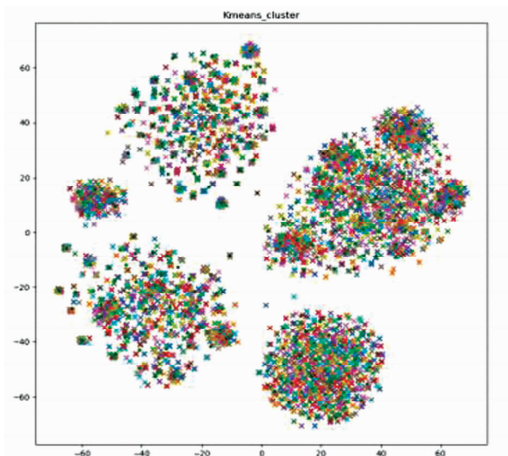


图 7 K-means 聚类结果图

的方法实现映射,但还有墙、天花板等,因为 IFC 标准下的实体表示和 CityGML 标准下的实体表示不同,前者为扫描体表达,后者为边界描述,因此还需要通过几何重构对这类实体进行进一步的映射。通过对提取出的“实体—关系—实体”三元组,进行遍历提取信息,得到实体的尺寸、坐标等,并根据这些信息进行几何重构,进一步实现从 IFC 标准的实体到 CityGML 标准的实体的映射和数据信息的补充,使匹配模型更加完善。

首先提取三元组中关系为 type 的集合。通过提取三元组信息得到的模型中墙这一类的具体实例的集合,如表 3 所示。

表 3 墙的实例

实体 1	关系	实体 2
IfcWallStandardCase_1339	type	IfcWallStandardCase
IfcWallStandardCase_4395	type	IfcWallStandardCase
IfcWallStandardCase_1509	type	IfcWallStandardCase

其次再根据三元组对墙的实例进一步深入扩展,获取到墙的尺寸、坐标等更详细的数据信息。如图 7 所示三元组中的数据关系:(a)获取墙的名字;(b)获取墙的高度;(c)获取墙的一层平面上四个点的坐标。

最终获取到了墙的实例的尺寸、一层平面四个点的坐标和高度,如表 4 所示 IfcWallStandardCase_133 的厚度为 200mm,高度为 8000,再根据四个点的坐标可知,该墙的尺寸长为 12400,高为 8000。根据墙的尺寸和坐标,可判断出墙的内表面,墙的外表面,其中墙的内表面映射到 CityGML 标准的 Interior-

WallSurface, 墙的外表面映射到 CityGML 标准中的 WallSurface。

表 4 IfcWallStandardCase_1339 厚度和坐标

厚度	200mm
高度	8000
坐标 a	(0, -100.00000000000007, 0)
坐标 b	(12400, -100.00000000000007, 0)
坐标 c	(12400, 99.99999999999993, 0)
坐标 d	(0, 99.99999999999993, 0)

通过对 IFC 标准的实体的数据信息的研究,发现其模型内部的实体坐标都以每一个实体自己的中心为原点,向外扩展。再通过对坐标的研究,发现如在墙这一实例中,从中心点扩展,向内墙方向的扩展为负,外墙方向的扩展为正。由此可以判断出,在表 4 中所列出的四个点 a,b,c,d 中,a,b 两点位于内墙,c,d 两点位于外墙。由此可分别根据 a,b 点和墙的尺寸、c,d 点和墙的尺寸来建立内墙和外墙两个面,并分别映射到 CityGML 标准中的 InteriorWallSurface 和 WallSurface 两个实体。同理可对应出天花板、地板等实体的内外表面。

3 在三维模型层面验证本体匹配模型

3.1 建立 IFC 标准的模型

用 Revit 建立一个 IFC 标准的模型并导出其 .ifc 文件,再通过转换得到 .owl 格式的文件。

(1)用 Revit 建立一个城市书房的三维建筑模型。并导出 IFC 文件。

(2)在导出 IFC 文件之后,利用 buildingSMART 官网中的 IFCConvertor,名字叫 UGent-Aalto IFC - to

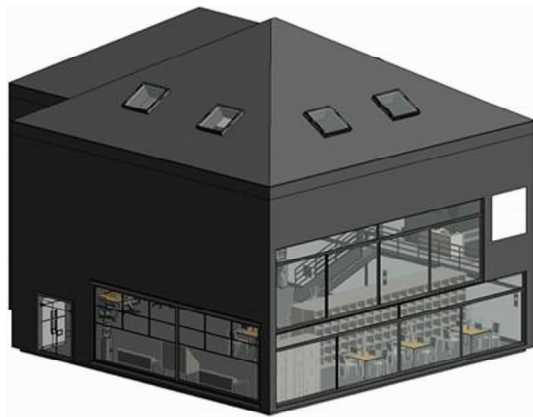


图 9 IFC 标准的三维模型

- RDF converter,并利用其中的 JAR 包将 IFC 文件转换为 ifcOWL 文件。

IFC 标准是利用 XML 数据格式描述数据内容,而 OWL 语言也是基于 XML 数据格式进行存储的。基于这一特性,可构建 IFC 到 OWL 的数据转换方式,将 BIM 模型的信息转换为可构建知识图谱的信息格式。Pauwels 等提出了一种 EXPRESS 到 OWL 的转换模式,该模式转换的 OWL 本体数据可实现数据的重用和查询、推理。根据该转换模式,实现对 IFC 中的实体类型和属性关系到 CityGML 标准的转换。

3.2 建立 CityGML 标准的模型

根据 1 中建立的实体匹配模型,将建立的 IFC 标准的模型对应映射出 CityGML 标准的数据信息,并根据这些数据建立一个 CityGML 标准的模型。

通过相同的方法解析 owl 文件,我们从中提取到了概念层文件 IFC4.0 所不具有的实体的详细信息,如表 5,表 6 所示,其中表 6 中通过字符串匹配、

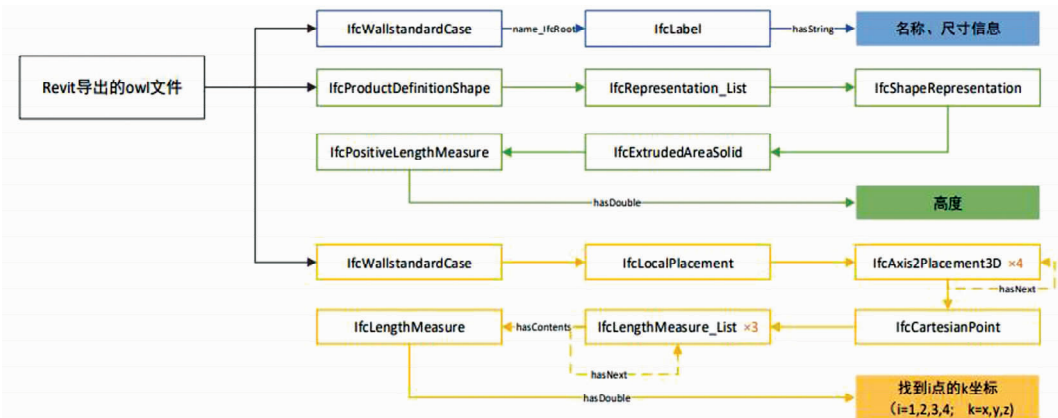


图 8 获取数据信息

表 5 三元组中实体详细信息

实体名称	IfcWindow_798
实体详细信息	IfcRelDefinesByProperties_805 ,relatedObjects_IfcRelDefines ,IfcWindow_798 IfcRelContainedInSpatialStructure_197097 ,relatedElements_ IfcRelContainedInSpatialStructure ,IfcWindow_798 IfcRelDefinesByType_197183 ,relatedObjects_IfcRelDefines ,IfcWindow_798 IfcRelFillsElement_197271 ,relatedBuildingElement_IfcRelFillsElement ,IfcWindow_798 IfcWindow_798 ,tag_IfcElement ,IfcIdentifier_114422 IfcWindow_798 ,overallWidth_IfcWindow ,IfcPositiveLengthMeasure_114424 IfcWindow_798 ,objectPlacement_IfcProduct ,IfcLocalPlacement_197275 IfcWindow_798 ,overallHeight_IfcWindow ,IfcPositiveLengthMeasure_114423 IfcWindow_798 ,ownerHistory_IfcRoot ,IfcOwnerHistory_41 IfcWindow_798 ,objectType_IfcObject ,IfcLabel_114171 IfcWindow_798 ,type ,IfcWindow IfcWindow_798 ,name_IfcRoot ,IfcLabel_114409 IfcRelAssociatesMaterial_197153 ,relatedObjects_IfcRelAssociates ,IfcWindow_798 IfcWindow_798 ,representation_IfcProduct ,IfcProductDefinitionShape_792

表 6 IFC 标准实体和 CityGML 标准实体对应关系

IFC 标准的实体	对应 CityGML 标准的实体
IfcWindow	Window
IfcDoor	Door
IfcWallStandardCase	InteriorWallSurface WallSurface
IfcSlab	FloorSurface
IfcRoof	RoofSurface
IfcCovering	CeilingSurface OuterCeilingSurface

知识表示学习和几何重构实现了 IFC 标准到 CityGML 标准的映射。

4 结论

本文提出了一种新的关于研究 IFC 与 CityGML 匹配融合的方法,并且通过上述实验,我们成功的将 IFC 标准的 owl 文件中的实体 Ifcwindow 与 CityGML 标准的 owl 文件的实体 window 聚为同一类,验证了知识表示学习的方法的可行性。通过对应内墙和外墙,验证了几何重构的可行性。

但实验也存在一些不足之处:比如,我们聚类的结果有局限性。通过分析实验步骤可知,我们在如下方面存在不足:

(1)处理三元组

在处理三元组的过程,我们并没有恰当的筛选

出所有合适的三元组,而是简单的筛选异常三元组和去重,冗余的三元组信息干扰了 TransE 向量化的结果。

(2)向量化模型选取

我们使用的是 Trans 家族的 TransE 算法,算法本身最适合的是一对一的三元组,我们的三元组可能在与算法结合上存在问题。

本文基于扩展语义匹配的 BIM 和 GIS 三维建筑数据融合研究,采用的知识表示学习的方法可以有有效的解决 IFC 和 CityGML 两个不同标准对同一实体的语义理解偏差问题,并取得了一定的成果。当然该方法在 TransE 参数的确定,三元组信息的筛选方面都需要更进一步的研究和改善。此外,在几何重构方面,我们还可继续进行进一步的研究,将 IFC 标准中实体自己的中心点原点坐标整合转换为一个原点,并根据此原点转换拓展出模型内部某一构件的坐标,最终实现将基于构件本身的坐标转换为基于整个模型的坐标,并通过 SWRL 规则来完成映射和对应。

参考文献

- [1] 武鹏飞,刘玉身,谭毅,李建锋. GIS 与 BIM 融合的研究进展与发展趋势[J]. 测绘与空间地理信息,2019,42(1): 1-6.
- [2] 李佩瑶. 从 BIM 实体模型自动提取多细节层次 GIS 表面模型的方法[D]. 西南交通大学,2017.

- [3] Shinji Yamamura, Liyang Fan, Yoshiyasu Suzuki. Assessment of Urban Energy Performance through Integration of BIM and GIS for Smart City Planning[J]. Procedia Engineering, 2017, 180.
- [4] Wu B, Zhang S. Integration of GIS and BIM for indoor geovisual analytics[J]. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 2016, 41: B2.
- [5] Yichuan Deng, Jack C. P. Cheng, Chimay Anumba. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison[J]. Automation in Construction, 2016, 67.
- [6] Javier Irizarry, Ebrahim P. Karan, Farzad Jalaei. Integrating BIM and GIS to improve the visual monitoring of construction supply chain management [J]. Automation in Construction, 2013, 31.
- [7] IFC to CityGML Transformation Framework for Geo-Analysis: A Water Utility Network Case [C]//4th International Workshop on 3D Geo-Information, 4 - 5 November 2009, Ghent, Belgium. 2009.
- [8] El - Mekawy M, Östman A, Shahzad K. Towards interoperating CityGML and IFC building models: a unified model based approach [M]//Advances in 3D geo-information sciences. Springer, Berlin, Heidelberg, 2011, 73-93.
- [9] 张正航, 钱育蓉, 行艳妮, 等. 基于 TransE 的表示学习方法研究综述 [J]. 计算机应用研究, 2021, 38 (3): 656-663.
- [10] 陈华钧, 张文, 陈名杨, 等. 一种可解释的可生成规则的知识图谱嵌入表示学习方法: .
- [11] Flexa Caio, Gomes Walisson, Moreira Igor, Alves Ronnie, Sales Claudomiro. 多边形坐标系: 使用几何 DR 和确定版本的 t - SNE 可视化高维数据 [J]. 具有应用程序的专家系统, 2021, 175.
- [12] Ding L, Zhou Y, Akinci B. Building Information Modeling (BIM) application framework: The process of expanding from 3D to computable nD [J]. Automation in Construction, 2014, 46 (oct.): 82-93.
- [13] 陈光, 薛梅, 刘金榜, 等. 一种市政道路 BIM 设计模型与三维 GIS 数据集成方法 [J]. 地理信息世界, 2018, v. 25; No. 129 (3): 90-94, 98.
- [14] Amirebrahimi S, Rajabifard A, Mendis P, et al. A framework for a microscale flood damage assessment and visualization for a building using BIM-GIS integration [J]. International Journal of Digital Earth, 2016, 9 (4): 363-386.
- [14] 张佳琪. BIM 与 GIS 数据融合方法研究 [D]. 长春工程学院, 2019.

BIM and GIS 3D Building Data Fusion Based on Extended Semantic Matching

Hu Yingting, Ma Jun, Shi Yu

(School of Shandong Normal University, School of Information Science and Engineering, Jinan 250300, China)

Abstract: At present, the research on the integration of BIM and GIS is mainly divided into three directions, namely data format conversion, data standard extension and ontology. This project chooses to carry out the entity integration of IFC and CityGML through the ontology mode. The triples of entities are extracted through semantic modification, and then the triples are represented as entity vectors by knowledge representation learning. The semantic information of entities is transformed into a low-dimensional vector space, and the matching fusion system of IFC and CityGML standards is constructed based on extended semantic matching. To solve the problem of semantic understanding bias of the same entity under IFC and CityGML, and to cluster by similarity calculation. In addition, the extracted triples were used to carry out geometric reconstruction using entity-relations - entity, and the entities of IFC standard were mapped to entities of CityGML standard. Finally, the established 3D model is used to verify and realize the transformation from BIM field to GIS field.

Key Words: BIM; GIS; IFC; CityGML; Knowledge Representation Learning; Geometry Reconstruction; Fusion