

# 顾及地理实体语义相似度的土地用途分区模型

牛继强<sup>1</sup> 徐 丰<sup>1</sup> 李卓凡<sup>1</sup> 洪晓峰<sup>2</sup>

<sup>1</sup> 信阳师范学院城市与环境科学学院,河南 信阳,464000

<sup>2</sup> 长江水利委员会长江科学院,湖北 武汉,430010

**摘 要:**结合土地用途分区工作的要求,提出了地理实体语义相似度测度模型;在对小图斑的模式进行划分的基础上,构建了顾及语义相似度的土地用途分区模型,并提出分区结果的面积平衡和最优化评价方法。以海南省昌江县海尾镇为例对模型进行验证,结果表明,该模型是一种科学、合理、高效的土地用途分区模型,该方法提供的土地用途分区方案不仅能够较好地反映认知特性,还能较好地概括土地利用的特征,可以为土地利用规划、土地利用调控和管理提供技术支持。

**关键词:**地理实体;语义相似度;语义距离;土地利用规划;土地用途分区

中图法分类号:P208

文献标志码:A

土地用途分区是土地利用总体规划的核心问题之一,我国目前在县、乡镇级土地利用总体规划中主要利用土地用途分区作为土地利用的调控和管理的重要依据<sup>[1,2]</sup>。为了能够自动地进行土地用途分区,国内外学者已经进行了大量研究。目前,土地用途分区方法主要体现在以下三个方面:①传统方法,如叠置分析法、主导因素法、多因素综合评判法等<sup>[3-8]</sup>,这些方法在解决具体的土地用途分区问题中有着一定的科学性和可操作性,但是存在效率低下、受人为作用的影响,不能满足海量数据处理的需求等问题;②基于模型的方法,如生态位适宜度模型、聚类模型等<sup>[8,9]</sup>,在分区的过程中注重其他因素的影响,取得较好的效果;③基于多目标的优化算法,如基于多目标的微粒群、禁忌搜索算法、CA、模拟退火和人工免疫系统等的最优化方法<sup>[10-13]</sup>,这些算法大多存在依赖参数,并且在执行效率和多目标处理方面还有待改进。土地用途分区已成为目前土地利用规划研究的热点问题之一。如何高效地获取土地用途分区方案是未来的重要研究内容。但已有的土地用途分区模型中没有考虑图斑的专题属性问题,有的虽有考虑,但是以主观判断为主。因此,在实现土地用途分区中,将土地利用图斑进行合并是土地用途分

区中最基本的问题。然而如何判断图斑的合并是一个复杂的决策。本文提出兼顾空间特征和语义特征的综合距离计算方法,并构建顾及语义相似度的土地用途分区模型,选择典型地区开展应用研究,以期获得一种高效的土地用途分区方法。

## 1 地理实体语义相似度测度模型

在地理信息科学中,已有用本体的概念属性和语义距离来测度地理实体的语义相似度的相关研究<sup>[14,15]</sup>,但是考虑的因素比较单一,不能全面地表达地理实体之间的语义相似度。本文结合概念的属性、语义距离等影响语义相似度的因素,提出了一种基于本体结构的地理实体语义相似度测度模型,同时将基于语义距离的方法与基于属性的方法相结合,从而更加全面地量化本体结构中实体之间的语义相似度。

### 1.1 概念的属性

在本体结构中,概念的属性是决定语义相似度的重要因素<sup>[14]</sup>。当两个概念拥有的相同属性越多,表明这两个概念间的语义相似度越大。概念属性的相似度计算公式为:

$$\text{Sim}^{\text{Att}}(S, F) = \frac{\text{Count}(\text{att}(S) \cap \text{att}(F))}{\text{Count}(\text{att}(S) \cup \text{att}(F))} \quad (1)$$

收稿日期:2013-11-06

项目来源:国家自然科学基金资助项目(41201387,41001219);河南省高等学校青年骨干教师计划资助项目(2012-GGJS-127);数字制图与国土信息应用工程国家测绘地理信息局重点实验室开放研究基金资助项目(GCwd201207)。

第一作者:牛继强,博士,副教授,现从事空间分析、空间数据挖掘等领域的研究。E-mail:niujiqiang@whu.edu.cn

其中,  $att(S)$  表示实体  $S$  属性的集合;  $att(F)$  表示实体  $F$  属性的集合;  $Count()$  表示统计出的属性个数。若地理实体的某种相应的性质不存在时, 则不用表示  $S, F$  在该性质上的相似度。

### 1.2 语义距离

语义距离是指本体结构中任意两个概念节点之间的最短路径长度。传统的语义距离的计算大都使用 Dijkstra 算法, 没有考虑有向边权重的影响<sup>[16]</sup>。本文对其进行了改进, 通过概念的关系、概念深度和概念密度三种权重因子把有向边的差异程度量化地表达出来, 公式如下:

$$Weight(S, F) = \alpha * Weight^{Type}(S, F) + \beta * Weight^{Depth}(S, F) + \gamma * Weight^{Density}(S, F) \quad (2)$$

其中,  $Weight(S, F)$  为有向边权重;  $Weight^{Type}(S, F)$  是概念的不同类型对应的语义距离的权重;  $Weight^{Depth}(S, F)$  是概念结点  $S, F$  之间的语义距离权重;  $Weight^{Density}(S, F)$  表示概念节点  $S, F$  之间有向边的权重关系;  $\alpha, \beta, \gamma$  是可调节因子, 且  $\alpha + \beta + \gamma = 1$ 。

由于有向边权重的大小与概念节点间的距离成反比, 因此权重大小与有向边语义距离的关系如下:

$$Dist(S, F) = \frac{1}{Weight(S, F)} \quad (3)$$

其中,  $Dist(S, F)$  为概念节点间的距离。

综上所述, 可得到改进的语义距离相似度计算方法如下:

$$Sim^{Dist}(S, F) = \frac{\mu}{Dist(S, F) + \mu} \quad (4)$$

其中,  $Sim^{Dist}(S, F)$  为语义距离相似度;  $\mu$  为可调节因子, 且为大于 0 的实数。

### 1.3 语义相似度计算

将基于语义距离的方法与基于属性的方法相结合, 可以得到地理实体语义相似度计算公式:

$$Sim(S, F) = \omega_1 * Sim^{Att}(S, F) + \omega_2 * Sim^{Dist}(S, F) \quad (5)$$

其中,  $Sim(S, F)$  为地理实体的语义相似度;  $\omega_1 + \omega_2 = 1$ , 且  $\omega_1, \omega_2$  大于 0。

## 2 土地用途分区模型的构建

### 2.1 小图斑模式的划分

土地利用方式多样, 造成土地利用空间布局的零散, 因此按照统一的土地用途管制规则划定土地用途区<sup>[2]</sup>, 是土地利用调控和管理的现实要求。在县、乡级规划编制中, 依据规划目标和土地

用途分区要求, 在与土地利用结构和布局调整方案相衔接的基础上, 拟定土地用途分区方案<sup>[2]</sup>, 其主要手段就是对邻近图斑进行合并, 具体就是将小图斑合并到邻近的大图斑中, 以形成利于管理的土地用途区。图斑的合并主要考虑到相邻图斑的邻近关系。在邻近关系判断时, 仅仅通过拓扑关系判断邻近是不完备的, 应综合考虑图斑的拓扑关系、几何关系和语义关系图斑的邻近性<sup>[15]</sup>。设原始图斑群集合为  $T = \{T_i, i = 1, 2, \dots, N\}$ ,  $N$  为图斑群内图斑的个数; 第  $k$  类土地利用类型图斑的最小上图面积记为  $\epsilon_k (k = 1, 2, \dots, M)$ ,  $M$  为  $T$  中土地利用类型的个数。以  $\epsilon_k$  为条件, 在  $T$  中检索面积小于  $\epsilon_k$  的所有图斑(在土地用途分区中称为小图斑), 组成小图斑集合  $R = \{R_j, j = 1, 2, \dots, L\}$ 。通过分析, 小图斑在土地规划图中存在孤立、线状邻接和聚集三种状态; 在与大图斑的拓扑关系中又分别有位于大图斑的内部和边界两种形式, 因此土地用途分区中的小图斑存在如图 1 所示的六种模式。图 1 中, 1~6 表示不同类型的小图斑, 灰色的图斑是面积小于  $\epsilon_k$  的小图斑。

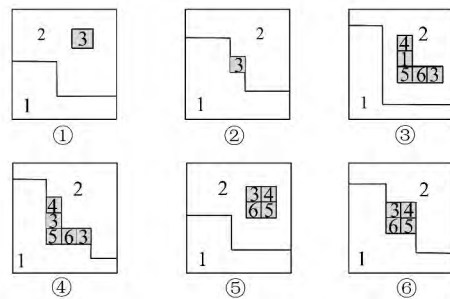


图 1 土地用途分区中小图斑存在的模式

Fig.1 Patterns of Small Polygons in Land Use Zoning

### 2.2 土地用途分区策略

土地用途分区有自上而下和自下而上两种类型合并方法, 本文提出的土地用途分区模型采用自下而上的分区方法, 并在土地分区过程中结合土地利用结构和布局调整方案进行分区。分区前, 应根据文献<sup>[17]</sup>中的要求针对不同的研究尺度定义不同用地类型的小图斑。具体方法如下: 首先将土地利用现状类型转换为土地规划用途分类中的类型, 然后根据表 1 的策略进行分区, 其中专项规划是划定土地用途分区的基础, 如某大图斑的线状类型可能是水田、旱地或其他的利用类型, 但是该图斑如果在专项规划中划定为基本农田, 则将其划分到基本农田保护区。根据这一方法, 对区域内的每个大图斑都给出了土地用途分区的类型。然后对小图斑进行分析, 根据土地利用现状类型的语义相似度和空间关系, 将小图斑

合并进大图斑,以大图斑的土地用途分区类型确定合并后图斑的类型。

表1 土地用途分区的划定方法  
Tab.1 Partition Method of Land Use Zoning

土地利用类型	专项规划用地类型	土地用途分区类型
耕地、园地	基本农田保护区	基本农田保护区
耕地、园地、其他农用地	土地开发整理规划	一般农地区
林地	林业规划	林业用地区
草地	草场规划	牧业用地区
城市、建制镇	城市、城镇规划	城镇建设用地区
村庄	村镇规划	村镇建设用地区
采矿用地、机场用地、港口码头用地	产业发展规划、工业园区规划	独立工矿区
风景名胜设施用地	旅游规划	风景旅游用地区
耕地、园地、其他农用地、林地、草地	地质灾害防治规划、生态环境规划	生态环境安全控制区
耕地、园地、其他农用地、林地、草地	自然保护区规划	自然和文化遗产保护区

对不同模式的小图斑,采用不同的方法进行合并。对于第①类模式,可以直接将小图斑合并到邻接的大图斑中。对于第②类模式,需要利用小图斑与相邻的大图斑的语义邻近分析,将其合并到语义邻近的大图斑中。

根据文献[2],对于第③、④种模式,要禁止小图斑合并成为新的土地用途区。其中第③类模式根据小图斑与大图斑语义相似度的高低依次进行合并;第④类模式根据小图斑与两个大图斑语义相似度的高低依次进行合并。

对于第⑤类模式,首先要对图斑群进行分析,分析2、3、4、5、6图斑之间的语义相似度。通过分析,存在三种情况:第一,小图斑之间的语义相似度较高,与大图斑之间的语义相似度也较高,可以将小图斑与大图斑进行合并;第二,小图斑之间的语义相似度较高,与大图斑之间的语义相似度较低,可以将小图斑合并,并判断是否大于 $\epsilon_k$ ,否则转化为①的问题求解;第三,一部分小图斑的语义相似度较高,其他小图斑与大图斑的语义相似度较高,先将语义相似度较高的大图斑和小图斑合并,然后将其他的小图斑进行合并,并判断是否大于 $\epsilon_k$ ,转化为①的问题求解。

对于第⑥类模式,首先要对图斑群进行分析,分析相互邻近图斑之间的语义相似度,存在两种情况:第一,小图斑之间的语义相似度较高,与大图斑之间的语义相似度较低,可以将小图斑合并,合并后的图斑如果大于 $\epsilon_k$ ,则形成一个独立的分区,否则转化为②的问题求解;第二,部分小图斑与邻近的大图斑的语义相似度较高,与其他小图

斑的语义相似度较低,先将语义相似度较高的大图斑和小图斑合并,其他小图斑合并,然后判断是否大于 $\epsilon_k$ ,否则根据其所属模式转化为①、②、④的问题求解。

### 2.3 土地用途分区算法

在描述顾及语义相似度的土地用途分区算法前做两个约定:①江、河、湖泊、水库、居民点以外的铁路、公路、管道运输用地、水利设施中的沟渠和水工建筑物用地不参与分区[2];②空间邻接。为了避免分区方案提供的分区形状不符合文献[2]的要求,采用文献[18]空间邻近度的计算方法,考察相邻的两个图斑中面积较小的一个共享边长与其周长的比值,如果该值小于0.2,则认为两个图斑不邻接。

在土地用途分区过程中,很多是将其转化为最优化问题进行求解,为了简化问题的求解方法,提高土地利用图斑合并的效率,具体算法如下。

1)在当前图斑群 $T$ 中提取面积小于 $\epsilon_k$ 的小图斑集合 $R$ ,对 $R$ 中的小图斑按照面积从小到大排序。

2)以面积最小的 $R_j$ 为当前图斑,对其邻域进行搜索。如果当前图斑的邻域内不存在面积小于 $\epsilon_k$ 的图斑,则判断图斑的模式。

a. 如果属于模式①,则将小图斑直接合并到大图斑,合并后的分区类型为大图斑的类型,转到1)。

b. 如果属于模式②,则计算小图斑与其邻接图斑的语义相似度,将小图斑合并到语义相似度较高的大图斑中,合并后的分区类型为大图斑的类型,转到1)。如果当前图斑的邻域存在面积小于 $\epsilon_k$ 的图斑,则继续搜索邻域内的小图斑,并将当前小图斑 $R_j$ 与其邻域的小图斑形成新的小图斑集合 $R'$ ,并判断图斑的模式。

c. 如果属于模式③,则分别计算小图斑与大图斑的语义邻近度,并按其大小将小图斑依次合并到大图斑,合并后的分区类型为大图斑的类型,转到1)。

d. 如果属于模式④,则计算每个小图斑与其邻接的大图斑的语义相似度,并按其大小将小图斑合并到语义相似度较高的大图斑中,合并后的分区类型为大图斑的类型,转到1)。

e. 如果属于模式⑤、⑥,则计算两两相邻的图斑的语义相似度,并对其从高到低进行排序,取语义相似度最高的两个图斑进行合并,合并后的分区类型取面积大的为新图斑的类型,判断新图斑的面积是否大于 $\epsilon_k$ ;生成新的小图斑集合 $R''$ ,判断 $R''$ 的模式,如果属于模式⑤、⑥,则继续采用

该方法进行合并,否则根据其模式分别转到步骤 a、b、c 或 d。

3) 判断小图斑集合 R 是否还有小图斑,若有,则返回步骤 1),否则转到步骤 4)。

4) 调整合并后的土地用途区,使合并后的变化值达到最小。

### 2.4 土地用途分区结果的评价

土地用途分区的过程可以认为是在一定约束条件下的最优解的获取过程。土地用途区面积约束是最基础的,理想情况下,图斑的转换可以相互抵消,但实际情况下,区域内的主导用途区合并的小图斑会较多,致使面积增加较多,在上文中已经给出解决办法,通过调整参数,或对合并的小图斑进行调整,使各土地用途区符合约束条件。

文献[19]给出了制图综合中图斑合并的约束条件,本文通过简化约束条件,提出土地用途分区的约束条件,用语义相似度给出最优化的处理方法。设土地用途分区过程中合并了 z 个小图斑,每个图斑的原始类型为  $S_{pk}$  ( $p=1,2,\dots,z;k=1,2,\dots,M,M$  为土地用途分区的类型数),每个图斑的面积为  $A_p$  ( $p=1,2,\dots,z$ ),这 z 个小图斑的类型集合为 SS,则土地用途分区的最优化控制目标为<sup>[20,21]</sup>:

$$\min \sum_{p=1}^z |S_{pk} - S_{pm}| \cdot A_p / \sum_{p=1}^z A_p \quad (6)$$

其中,  $S_{pm}$  表示第 p 个小图斑合并后所属的土地

利用类型;  $|S_{pk} - S_{pm}|$  表示第 p 个小图斑所属的原始土地利用类型与合并后所属的土地利用类型之间的语义相似度的差值。约束条件是  $SS_{pk} \in SS, SS_{pm} \in SS, SS_{pk} \neq SS_{pm}, 1 \leq p \leq z, 1 \leq m \leq M$ , 其中,  $SS_{pk}$  是 z 个小图斑原始土地利用类型构成的集合;  $SS_{pm}$  表示 z 个小图斑合并后所属土地利用类型构成的集合。理论上,第 p 个图斑的原始类型  $S_{pk}$  可以转换为集合 SS 中的任何一类,土地用途分区过程中已记录了该图斑可转换的图斑类型集合,因此根据式(6)可以评价出计算的结果是否为最优解,如果不是,可以采用调整图斑的最小上图面积指标的方法重新进行分区。

## 3 实验研究

### 3.1 土地利用类型的语义相似度实验

根据本文提出的地理实体语义相似度测度模型,对现行的标准土地利用分类中的实体转换为土地规划用途分类中的类型,并将其作为本体对象进行语义相似度测算实验。根据各种影响因素对语义相似度的影响程度的差异,通过实验分析及验证,将模型中的各个调节因子分别取值如下:  $\alpha=0.52, \beta=0.27, \gamma=0.21, \mu=5, \omega_1=0.8, \omega_2=0.8$ 。计算的各土地利用类型之间的语义相似度值(部分)见表 2。

表 2 各土地利用类型之间的语义相似度  
Tab. 2 Semantic Similarity Between Land Use Types

	耕地	园地	林地	牧草地	其他农用地	城镇用地	农村居民点用地	采矿用地	其他独立建设用地	港口码头用地	风景名胜设施用地	特殊用地	自然保留地
耕地	1	0.437	0.437	0.437	0.445	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.107
园地	-	1	0.432	0.432	0.440	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.100
林地	-	-	1	0.432	0.440	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.100
牧草地	-	-	-	1	0.440	0.267	0.267	0.267	0.267	0.267	0.267	0.267	0.100
其他农用地	-	-	-	-	1	0.282	0.282	0.282	0.282	0.282	0.282	0.282	0.112
城镇用地	-	-	-	-	-	1	0.446	0.446	0.446	0.439	0.439	0.439	0.107
农村居民点用地	-	-	-	-	-	-	1	0.446	0.446	0.439	0.439	0.439	0.107
采矿用地	-	-	-	-	-	-	-	1	0.446	0.439	0.439	0.439	0.107
其他独立建设用地	-	-	-	-	-	-	-	-	1	0.439	0.439	0.439	0.107
港口码头用地	-	-	-	-	-	-	-	-	-	1	0.439	0.439	0.107
风景名胜设施用地	-	-	-	-	-	-	-	-	-	-	1	0.283	0.107
特殊用地	-	-	-	-	-	-	-	-	-	-	-	1	0.107
自然保留地	-	-	-	-	-	-	-	-	-	-	-	-	1

### 3.2 土地用途分区实验结果与分析

采用本文提出的模型进行土地用途分区实验。该实验采用 Microsoft Visual. NET 2008 作为开发平台,在 Arc Engine 组件基础上二次开发定制 GIS 功能;采用编写代码实现本文提出的土地用途分区模型。以海南省昌江县海尾镇 1:1 万

土地利用现状图(见图 2)为实验数据,根据本文提出的土地用途分区策略和模型,对实验区进行实例研究。

将海南省昌江县海尾镇 4 412 个土地利用现状图斑进行图斑转换,转换为土地规划用途分类中的类型。根据不同类型的最小上图面积要求,

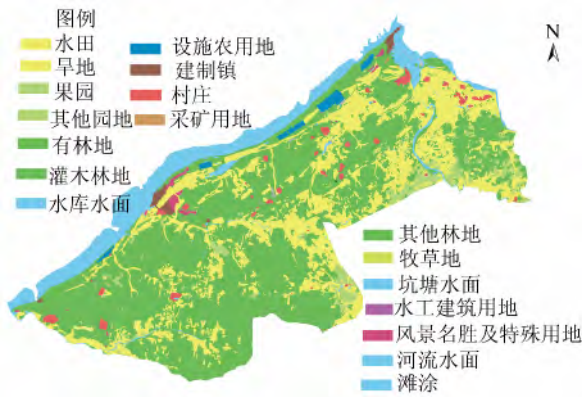


图2 研究区土地利用现状图

Fig. 2 Map of Land Use in Study Area

需要对 1 345 个小图斑进行合并,然后根据本文构建的模型对其进行分区,共形成 441 个土地用途区,实验结果如图 3 和表 3 所示。根据式(6),在实验过程中记录了土地用途分区的最优化控制目标,其结果为 10.632,平均语义相似度差值为 0.007 91。由于每次合并的小图斑的语义相似度较高、面积较小,因此其变化趋势呈线性增加。

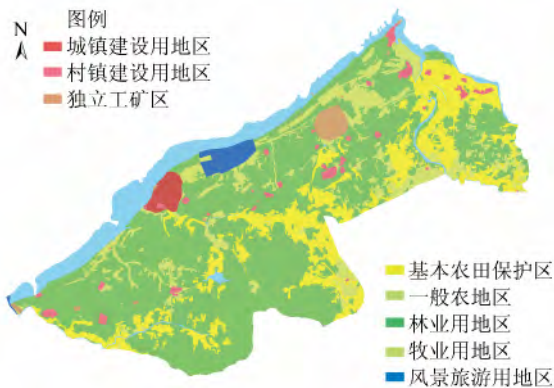


图3 顾及地理实体语义相似度的土地用途分区图

Fig. 3 Map of Land Use Zoning Based on Semantic Similarity of Geographic Entity

海尾镇土地利用总体规划(2005~2020年)中的土地利用分区是在GIS环境下使用叠置分析并通过局部调整完成的土地用途分区(图4)。将两种方案进行对比分析,本文的方案中各类用地类型的面积变化不大,用地类型间的面积比例基本保持平衡,图斑数量有大幅的减少,分区结果较好地概括了土地利用特征和反映了人类的认知特性,符合相关标准的要求,是较为理想的土地用途分区方案。

本文的土地用途分区模型主要通过枚举法计算最优解,因此运算时间仅与小图斑的个数有关。从实验过程来看,在县乡级的土地用途分区中合并的图斑是有限的,因此本文的算法设计能够保

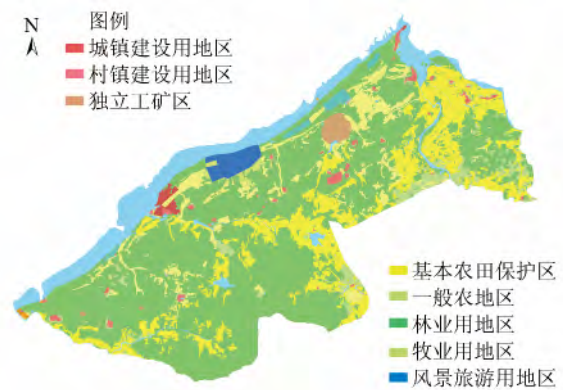


图4 基于叠置分区的土地用途分区图

Fig. 4 Map of Land Use Zoning Based on the Overlay

证较高的运算效率。从实验过程来看,在保证得到理想的土地用途分区的前提下,本文提出的模型避免了部分智能优化算法存在过拟、容易陷入局部最优等方面的问题。

在本文实验结果中,城镇建设用地区、村镇建设用地区、独立建设用地区和风景旅游用地区比土地利用总体规划中的土地用途区面积较大,主要原因是本文方案中包含了有条件建设区,符合文献[2]中关于有条件建设区布局的要求。通过对分区后的评价,本实验的结果是较为理想的土地用途分区方案。

表3 土地用途区数量与面积表

Tab. 3 Quantity and Area of Land Use Zoning

分区类型	规划面积/m <sup>2</sup>	分区面积/m <sup>2</sup>	图斑数/个
城镇建设用地区	102.8	184.6	1
村镇建设用地区	335.5	438.8	35
独立建设用地区	209.8	246.9	13
基本农田保护区	3 669.0	3 693.1	105
一般农地区	2 798.6	2 405.8	219
林业用地区	12 109.7	11 989.1	64
牧业用地区	6.5	4.8	2
风景旅游用地区	10.9	249.0	2
河流及其他	2 017.1	2 047.8	27
合计	21 259.9	21 259.9	441

### 4 结 语

土地用途分区需要同时兼顾空间邻接、几何形状与语义相似度等因素,其核心是综合距离的判别。本文提出了地理实体的语义相似度测度模型,并在图斑的模式进行划分的基础上,构建了顾及语义相似度的土地用途分区模型。从实例研究来看,将语义相似度应用于土地用途分区,通过科学设定参数,可以在一定程度上考虑空间邻近、几何形状等因素对次要图斑进行合并,获得土地用途分区方案。实验表明,该模型所获得的土地用

途分区方案较好地反映了人类的认知特性,概括了土地利用的特征,可以满足土地利用规划中的需求,并且分区效率有所提升。本模型主要考虑地理实体的语义相似度,仅给出了空间邻接和集合形状等方面的一些限定条件,未来的研究还可以通过结合图斑的空间邻接和分区后的几何形状等问题进一步对土地利用分区模型进行研究。

### 参 考 文 献

- [1] Niu Jiqiang, Liu Minghua, Xu Feng. A System of Land Use Zoning Based on Spatial Data Mining[J]. *Journal of Xinyang Normal University (Natural Science Edition)*, 2012, 25(3): 382-386(牛继强, 刘明华, 徐丰. 基于空间数据挖掘的土地用途分区系统[J]. 信阳师范学院学报(自然科学版), 2012, 25(3):382-386)
- [2] Ministry of Land and Resources of the Peoples Republic of China. TD/T 1205-2010. Standards for Overall Planning of Land Utilization at the Township (Town) Level[S]. Beijing: China Standard Press, 2010(中华人民共和国国土资源部. TD/T 1205-2010 乡(镇)土地利用总体规划编制规程[S]. 北京: 中国标准出版社, 2010)
- [3] Cheng Ye, Wang Jing, Meng Fanhua. Research on Land Use Zoning and Regulation[M]. Beijing: Geological Press, 2003:6-73(程烨, 王静, 孟繁华. 土地利用分区管制研究[M]. 北京:地质出版社, 2003: 6-73)
- [4] Lin F T. GIS-based Information Flow in a Land-use Zoning Review Process[J]. *Landscape and Urban Planning*, 2000, 52(1): 21-32
- [5] Pan Jinghu, Liu Yang, Shi Peiji. Research on Overall Planning of Land Use Division in Tianshui Based on PCA and GIS[J]. *Soils*, 2011, 43(4): 637-643(潘竟虎, 刘扬, 石培基. 基于主成分分析和GIS的统筹市域土地利用分区研究[J]. 土壤, 2011, 43(4):637-643)
- [6] Cong Mingzhu, Ou Xiangjun, Zhao Qing, et al. Division of Land Use Degree in Jiangsu Province Based on Principal Component Analysis[J]. *Geographical Research*, 2008, 27(3): 574-582(丛明珠, 欧向军, 赵清, 等. 基于主成分分析法的江苏省土地利用综合分区研究[J]. 地理研究, 2008, 27(3): 574-582)
- [7] Meng Lina, Zheng Xinqi, Zhao Lu, et al. Land-use Functional Regionalization Based on Niche-fitness Model[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2011, 27(3): 282-287(蒙莉娜, 郑新奇, 赵璐, 等. 基于生态位适宜度模型的土地利用功能分区[J]. 农业工程学报, 2011, 27(3):282-287)
- [8] Zhao Rongqin, Huang Xianjin, Zhong Taiyang, et al. Application of Clustering Analysis to Land Use Zoning of Coastal Region in Jiangsu Province[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2010, 26(6): 310-314(赵荣钦, 黄贤金, 钟太洋, 等. 聚类分析在江苏沿海地区土地利用分区中的应用[J]. 农业工程学报, 2010, 26(6): 310-314)
- [9] Wang Hanhua, Liu Yanfang. Optimal Allocation of Land Resources Based on MOP-CA[J]. *Geomatics and Information Science of Wuhan University*, 2009, 34(2): 174-177(王汉花, 刘艳芳. 基于MOP-CA整合模型的土地利用优化研究[J]. 武汉大学学报·信息科学版, 2009, 34(2):174-177)
- [10] Liu Yaolin, Xia Yin, Liu Dianfeng, et al. Optimization of Land Use Zoning Based on Goal Programming and Simulated Annealing[J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(7): 762-765(刘耀林, 夏寅, 刘殿峰, 等. 基于目标规划与模拟退火算法的土地利用分区优化方法[J]. 武汉大学学报·信息科学版, 2012, 37(7):762-765)
- [11] Wang Hua, Liu Yaolin, Ji Yingli. Land Use Zoning Model Based on Multi-objective Particle Swarm Optimization Algorithm[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, 28(12): 237-244(王华, 刘耀林, 姬盈利. 基于多目标微粒群优化算法的土地利用分区模型[J]. 农业工程学报, 2012, 28(12):237-244)
- [12] Liu Y L, Liu D F, Liu Y F, et al. Rural Land Use Spatial Allocation in the Semiarid Loess Hilly Area in China: Using a Particle Swarm Optimization Model Equipped with Multi-objective Optimization Techniques [J]. *Science China Earth Science*, 2012, 55: 1 166-1 177
- [13] Liu Yaolin, He Jianhua, Liu Dianfeng. A Land-use Automatic Partitioning Method[P]. China Patent: CN101877034A, 2010-11-03(刘耀林, 何建华, 刘殿峰. 一种土地利用自动分区方法[P]. 中国专利: CN101877034A, 2010-11-03)
- [14] Liu Ziyu, Huang Lei. Research on Concept Semantic Similarity Computation Based on Domain Ontology Model[J]. *Journal of the China Railway Society*, 2011, 33(1): 52-57(刘紫玉, 黄磊. 基于领域本体模型的概念语义相似度计算研究[J]. 铁道学报, 2011, 33(1):52-57)
- [15] Liu Yaolin, Li Hongmei, Yang Chunhui. Ontology Based Land Use Data Generalization[J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(8): 883-886(刘耀林, 李红梅, 杨淳惠. 基于本体的土地利用数据综合研究[J]. 武汉大学学报·信息科学版, 2010, 35(8):883-886)

- [16] Zhang Lanfang. Natural Language Semantic Similarity Algorithm Based on Ontology [J]. *Journal of Guilin University of Technology*, 2012, 32(2): 253-258 (张兰芳. 一种基于本体的自然语言语义相似度算法[N]. 桂林理工大学学报, 2012, 32(2): 253-258)
- [17] The Peoples Republic of China. TD/T 1205-2010. Standards for Land Use Condition Classification [S]. Beijing: China Standard Press, 2007 (中华人民共和国. GBT21010-2007 土地利用现状分类[S]. 北京: 中国标准出版社, 2007)
- [18] Ai Tinghua, Yang Fan, Li Jingzhong. Land-use Data Generalization for the Database Construction of the Second Land Resource Survey [J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(8): 887-891 (艾廷华, 杨帆, 李精忠. 第二次土地资源调查数据建库中的土地利用图综合缩编[J]. 武汉大学学报·信息科学版, 2010, 35(8): 887-891)
- [19] Niu Jiqiang. Spatial Data Mining for Land Use Zoning [D]. Wuhan: Wuhan University, 2010: 88-107 (牛继强. 面向土地用途分区的空间数据挖掘[D]. 武汉: 武汉大学, 2010: 88-107)
- [20] Guo Qingsheng, Wang Xiaoyan, Liu Jiping. Progressive Combination of Polygon Groups [J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(2): 220-232 (郭庆胜, 王晓妍, 刘纪平. 图斑群合并的渐进式方法研究[J]. 武汉大学学报·信息科学版, 2012, 37(2): 220-232)
- [21] Cheng T, Li Zhilin. Towards Quantitative Measures for Semantic Quality of Polygon Generalization [J]. *Cartographica*, 2006(2): 135-147

## Land Use Zoning Model Considering Semantic Similarity of Geographic Entity

NIU Jiqiang<sup>1</sup> XU Feng<sup>1</sup> LI Zhuofan<sup>1</sup> HONG Xiaofeng<sup>2</sup>

<sup>1</sup> School of Urban and Environmental Science, Xinyang Normal University, Xinyang 464000, China

<sup>2</sup> Changjiang River Scientific Research Institute, Changjiang Water Resources Commission, Wuhan 430010, China

**Abstract:** In terms of land use zoning, requirements, a semantic similarity measurement model of geographic entity is proposed. Methods for area balance and optimization evaluation are put forward for land use zoning partition results based on a study of the patterns of small polygons; a model of land use zoning was constructed considering semantic similarity. These models are verified in a case study of land use zoning in Haiwei, Changjiang county, Hainan province, China. The results show that the proposed methods provide a reasonable and efficient land use zoning model. Land use zoning schemes produced by these models, not only can reflect human cognitive features but also can summarize land use characteristics, and technically support land use planning and land use regulation and management.

**Key words:** geographic entity; semantic similarity; semantic distance; land use planning; land use zoning

**First author:** NIU Jiqiang, PhD, associate professor, specializes in spatial analysis & spatial data mining. E-mail: niujiqiang@whu.edu.cn

**Foundation support:** The National Natural Science Foundation of China, Nos. 41201387, 41001219; Youth Core Teachers Funding Scheme of University in Henan Province, No. 2012-GGJS-127; Open Research Fund Program of Key Laboratory of Digital Mapping and Land Information Application Engineering, National Administration of Surveying, Mapping and Geoinformation, No. GCwd201207.