

基于空间陈述的定位及不确定性研究

张毅¹, 邬阳², 高勇¹, 刘瑜^{1*}

(1. 北京大学遥感与地理信息系统研究所, 北京 100871; 2. 北京建筑工程学院测绘与城市空间信息学院, 北京 100044)

摘要: 在地理空间知识的表达中, 通常以文本方式描述位置。除了常见的地名以及地址外, 许多陈述根据与参照地物的空间关系来表达一个目标地物的位置。由于一个空间陈述只是粗略描述了目标对象的位置, 因此具有不确定性。在基于空间陈述的定位问题中, 不确定性包括4个层次, 分别对应于陈述、参照对象、空间关系和目标对象。这4个层次的不确定性适合于不同的建模方式, 如概率方法、模糊集方法, 以及证据理论方法等。本文采用不确定性场的概念, 对点状目标地物的分布进行了探讨, 并利用贝叶斯定理证明了对于给定的空间陈述, 其不确定性场分布与做出该陈述概率, 以及相应空间关系的模糊性之间的联系, 该结论可用于指导不确定性场的建立。

关键词: 空间陈述; 空间定位; 参照对象; 目标对象; 不确定性

DOI: 10.3724/SP.J.1047.2013.00038

1 引言

随着人工智能(Artificial Intelligence, AI)与地理信息系统(Geographical Information System, GIS)的结合, 定性地理空间知识日益受到重视^[1-3]。基于场所的GIS被提出和研究^[4-5], 与该问题相关的技术包括地理本体^[6]、空间关系的表达和推理^[7-9]、数字地名辞典^[10-12]、以及地理信息检索^[13]等, 在地理空间知识中, 位置信息多通过文本形式定性表达。由于以下两个方面的原因, GIS需要处理自然语言表达的空间位置。在实践中, 许多领域数据库中的位置描述也是基于文本方式, 例如, 在美国的MaNIS项目中^[14], 包含了大约30万条关于物种观测点的记录。这些记录大多缺少精确的地理坐标, 而仅仅是粗略的文本描述。如果能够对文本形式表达的位置确定其地理坐标, 就可借助于GIS软件进行进一步的空间分析。

Goldberg^[15]总结了表达位置的文本串的形式, 包括街道地址、邮政编码、命名地理要素, 以及一段基于空间关系的描述, 这些形式的文本串可以统称为广义地名^[16]。本文将根据一段文本确定被描述

点位置的技术称为“空间定位(Spatial Positioning)”。通常, 空间定位处理的文本包括如下3类: (1)地名, 如“北京”、“黄河”等; (2)地址, 如“北京市海淀区颐和园路5号”, 广义的地址, 如电话号码和IP地址等; (3)空间断言及组合, 如“中关村附近”或“北京大学以东”。数字地名辞典可以用于处理第一类文本, 并且是后两类文本处理的基础。目前地址的空间定位, 即地理编码(Geo-coding)问题, 已有较多研究, 而文献^[17]和^[18]探讨了基于空间陈述的定位问题。本文研究了空间定位的不确定性问题, 并以概率方法, 提出了不确定性场的空间定位模型。

2 空间陈述的概念

让我们考察如下陈述: “北京大学附近的小餐馆”。该陈述包括了3个要素, 即参照对象(Reference Object, 即北京大学)、空间关系(Spatial Relationship, 即附近)、和目标对象(Target Object, 即小餐馆)。在空间定位问题中, 参照对象与目标对象通过空间关系建立联系。尽管在空间关系表达中, 任何地物都可以成为一个表述的参照对象。然

收稿日期: 2012-10-19; 修回日期: 2012-12-11.

基金项目: 国家自然科学基金项目(41171295、41271385)。

作者简介: 张毅(1971-), 男, 博士, 讲师, 主要研究方向为地理信息系统应用。E-mail: zy@pku.edu.cn

* 通讯作者: 刘瑜(1971-), 男, 博士, 副教授, 主要研究方向为地理信息科学、人类移动模式与空间行为。

E-mail: liuyu@urban.pku.edu.cn

而,从知识表达和传输的角度,参照地物通常要足够显著、稳定和广为人知。

在地理信息科学中,空间关系一直是一个研究热点^[7,19-21]。这些研究既包括了3类二元空间关系,即拓扑关系、方向关系和度量关系,也包括一些特殊的空间关系,如“沿着”^[22]、“在...之间”^[23]等。空间关系的研究通常从形式化和空间认知两个角度入手,以建立符合人的表达习惯的可计算模型。在人工智能领域,通常是建立一个联合完备且两两互斥的关系集合作为形式化的基础。例如,在RCC8模型中,分辨了8种基本面拓扑关系^[24]。然而,在空间定位描述中,一些关系较少被使用。例如,人们较少用“不在...内”表达一个目标地物的位置。考虑到实际表达,本文建立空间定位问题中的空间关系概念图(图1)。

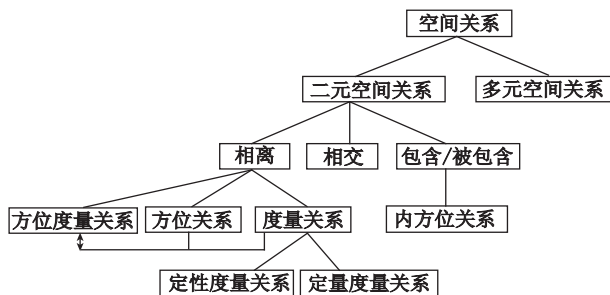


图1 空间陈述中所使用空间关系的层次化概念模型

Fig. 1 Hierarchical structure of spatial relations used in spatial assertions

根据图1,当两个对象相离时,可以采用方位关系、度量关系或者方位与度量关系的组合(如“北京大学以南500 m”)来对该关系进行精化;而当两个对象关系为包含或被包含时,可以采用内方位关系^[25]来进一步明确表达。值得指出的是:(1)包含和被包含关系可以在位置描述中直接被使用,而相离关系较少被单独使用来表达一个位置;(2)定性度量关系包括远、近等,然而人们通常只用“近”来描述目标对象的位置。

最后,目标对象是空间陈述的表达对象,也是空间定位研究中需要确定地理位置的对象。该对象通常具有维度和类型的约束。在常识性地理知识表达中,地理空间通常被抽象为二维空间^[26],因此,目标对象的维度可以为0、1和2,即点、线和面。目标对象的类型通常要显式给出,但有时可以不指定,如“在...附近见面”。当给出目标对象类型时,空间定位需要考虑地理背景知识,反之,则可以仅

考虑参照对象和空间关系。目标对象的维度有时可以被类型隐式蕴含,如河流类型通常意味着线状对象。此外,目标对象维度和空间关系也互相约束,例如,“沿着”、“平行”等关系同样意味着线状对象^[27]。当目标对象具有空间扩展时,空间定位问题需要对形状、尺寸的描述,然而目前对该问题的研究相对较少。因此,本研究将目标地物抽象为点而忽略其具体形状,它可以视为空间扩展对象的质心。

3 空间定位的不确定性

基于文本陈述表达空间位置,提供了一种有效的途径来表达和传输地理空间知识。然而,不确定性是人类智能的重要特征。在空间定位问题中,同样要处理不确定性。为了根据空间陈述建立目标对象的不确定性模型,首先需考察做出空间陈述的语境,本研究提出了关于语境的两个假设,归纳了与空间陈述相关的4个层次的不确定性,这两个假设和4个层次不确定性最终都影响到目标对象的空间分布。

3.1 空间定位中的两个假设

在空间定位问题中,目标点的位置由参照地物和空间关系所描述。为了进行空间定位,需要对陈述者描述目标点时的语境作出合理假设。根据参照对象和空间关系,有两个假设可以用于空间定位。假设一,陈述者认为目标对象的位置使得空间关系成立,假设二,陈述者认为目标对象的位置使得该空间关系最适合于描述目标地物的位置。

有两个方面,即空间关系的模糊性与多个参照地物情形下的可选性造成了上述两个假设的差异^[28]。在对目标对象陈述所采用的空间关系中,许多具有模糊性,如“近”,其隶属度与绝对距离的关系如图2(a)所示。假定两个对象A、B之间距离 d 对应的隶属度为 $0 \leq \mu(d) \leq 1$,则可以认为在一组观察者中,对“B在A附近”这个陈述给出肯定判断的比例是 $\mu(d)$ 。换言之,对于假定B与A的距离为 d ,并且 $\mu(d) > 0$,则存在部分观察者肯定“B在A的附近”。因此,根据假设一,当依据“在A附近”这样一个描述,B点是可能的候选点。

在实际中,观察者可以从多个空间关系,如“远”和“近”,中选择合适的。由于“远”关系也是模

糊的,可以假设其隶属度函数 μ_f 与近的隶属度函数 μ_c 相交于的 d_0 点。这样假定 B 与 A 的距离为 $d_1 > d_0$ 时,由于 $\mu_f(d_1) > \mu_c(d_1)$,观察者倾向于选择“远”来描述 B 点的位置。因此,基于假设二和“在 A 附近”, B 点应该被排除(图 2(b))。

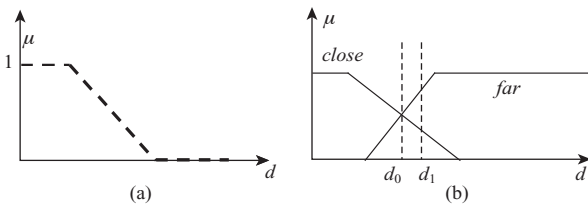


图 2 模糊空间关系对于空间位置表述的影响

Fig. 2 Vague spatial relations in the spatial assertion

除了空间关系的模糊性和多选择性外,参照对象也存在着多选择性^[29]。例如,“河南省信阳市在北京以南”是一个正确的空间陈述,然而,人们却很少用陈述来描述信阳市的地理位置,而是选择其他的参照地物,如郑州、武汉等更接近被描述点的城市。根据文献[30],可以根据参照地物构成的 Voronoi 图确定在描述某个目标对象时采用哪一个参照对象。如图 3 所示, R_1 到 R_8 是候选参照对象,并构成 Voronoi 图。为了描述 T_1 的位置,由于其在 R_1 对应的 Voronoi 多边形中,可以用 R_1 来描述 T_1 的位置,如“在 R_1 附近”或“在 R_1 东部”,而其他对象被用来作为参照对象的可能性则相对较低。

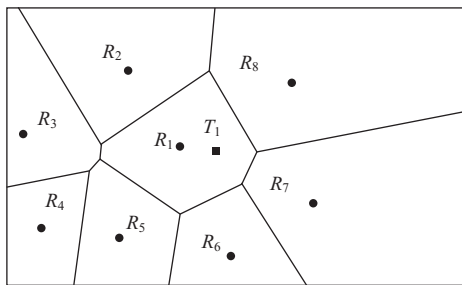


图 3 多参照地物对于空间位置表述的影响

Fig. 3 Multiple reference objects in generating a spatial assertion

鉴此,对于空间定位问题,假设一会导致更大的分布范围,即更大的不确定性。例如,根据空间陈述“在北京以南”,由假设一,信阳市也应该是目标对象可能的位置;而根据假设二,则可以排除信阳对应的位置。在空间定位中通常采用假设二,因为它更符合正常的知识加工过程;而假设一可以用于解释非定位目的的声明性空间知识。

3.2 空间陈述不确定性的 4 个层次

对于一个位置描述,其包含的不确定性主要包括以下 4 个层次:目标点的分布范围,空间关系的不确定性,参照对象的不确定性和陈述的不确定性^[31]。

首先,任何一个空间陈述只是表示了目标的一个分布范围,而不是准确的地理位置。由于在常识性地理空间知识中,地表空间被抽象为二维平面,因为该分布范围可以认为是 R^2 的一个子集。

其次,在一个空间陈述中,所采用的空间关系通常是不确定的。对于定性空间关系,如“近”、“北”等,具有含糊性(Vagueness)。空间关系的含糊性,使得目标对象的空间分布具有不均匀性,本文将给出一个证明,说明了空间关系含糊性对于目标对象分布的影响。其次,一个定量表达关系,如距离关系,由于测量误差以及记录精度的原因,也可以具有不确定性。例如,根据空间陈述“到点参照地物 A 的欧氏距离为 50km”,假定 A 点是确定的,则目标对象的分布为一个圆周,反之,如果度量距离是不确定的,则目标对象的分布为一个圆环带。除了上述不确定性外,对于一个定量度量关系,它可能指欧氏距离,也可能指两个对象之间的路径距离,这同样带来了目标对象的不确定性。

第三,参照对象也可能具有各种不确定性。一是由于参照对象是用自然语言表述的,这将带来二义性,因为存在重名的歧义地名。再是参照对象通常具有模糊性,如中关村^[32]。而后,即使参照对象有确定的边界,由于测量的原因,也可能存在误差。例如“离昌平 10km”,即使此空间陈述中“昌平”是指具有明确边界的北京市昌平区,也很难确定该距离是具体昌平区内哪一个点量算的,现有 GIS 软件通常是从地物边界起算,但是这显然不能满足空间定位的要求。文献[33]采用积分给出了模糊对象之间距离的计算方法,该方法同样可以用于起算点不确定情形下的度量关系处理。

最后,一个空间陈述本身可能有不确定性,这种不确定性来自于认知^[34],因此,可用证据理论^[35-36]对其进行建模。根据证据理论,可为每个空间陈述附加一个 mass 值,如 0.8,表明了表达者赋予该论断的信任。将证据理论引入空间定位问题,二维空间的每一个子集都对应了一个信任值,而未分配的信任则可以认为对应于环境。在空间定位问题中,环境是指根据一个应用(注意不是空间陈

述)确定的目标地物的可能最大分布范围。例如,在一个针对北京市的应用中,北京市域范围可以认为是环境。假定在该应用中有信任为 0.6 的空间陈述“在北京大学附近”,其剩余的信任应该被分配给环境,即整个北京市域。换言之,目标对象分布在北京市范围内的可能性是 0.4。从而,对于北京大学附近这样一个范围,目标地物分布于其中的下概率(或 Belief,可信度)为 0.6,而上概率(或 Plausibility,可能度)为 1.0。对于多个来源相互独立的关于同一目标对象的不确定陈述,可以采用 Dempster 组合法则进行证据组合,并计算其结果不确定性场。

4 基于不确定性场的定位模型

4.1 不确定性场

空间陈述的 4 个层次不确定性,最终都将影响目标对象的空间分布。假定目标对象被抽象为点,则根据给定空间陈述,可用一个二维概率密度函数 $p(x, y)$ 来表示目标点的空间分布。因为该函数以概率形式表示了目标对象的空间不确定性,并且具有二维场的形式,文献[18]称其为不确定性场(Uncertainty Field)。假定该函数的支持集(Support)为 F ,其拓扑维度可以为 0、1 或 2。对于一个确定的陈述,当维度为 2 时,有:

$$\iint_F p(x, y) dx dy = 1 \quad (1)$$

如果维度为 1 或 0 时,则分别需要用第一型线积分和加和来代替上式中的面积分。图 4 描述了不同空间陈述对应的不同支持集的形状。

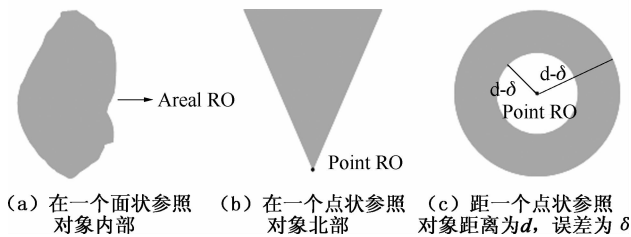


图4 不同空间陈述描述目标对象的不确定性场对应的支持集
Fig. 4 Supports of different uncertainty fields corresponding to various spatial predicates

- (a) inside an areal RO; (b) north of a point RO; and
(c) Euclidean distance d to a point RO, and the error is δ

为了在计算机中实现不确定性场,需要对其离

散化,文献[18]采用了栅格化的方式对其进行管理。对于每个栅格点 (i, j) ,其值为:

$$G(i, j) = \iint_C p(x, y) dx dy \quad (2)$$

其中, C 为栅格点 (i, j) 覆盖的区域。关于不确定性场更详细的描述见文献[18]。

当给定了某类空间陈述的不确定性场之后,就可以计算目标对象在二维空间中任意一个区域中的概率。在应用中,首先,给定一个概率值 P ,例如 0.95;然后,计算在此概率下的最小区域,称为 P 概率最小区域,它是 $p(x, y)$ 支持集的一个子集 R_P ,满足:

$$\begin{aligned} \iint_{R_P} p(x, y) dx dy &= P \\ \forall (x_1, y_1) \in R_P, \forall (x_2, y_2) \notin R_P, \\ p(x_1, y_1) &\geq p(x_2, y_2) \end{aligned} \quad (3)$$

特别的,当 $P=1.0$ 时, R_P 即为其支持集 F 。

4.2 不确定性场推导

空间陈述的定位问题实质是:根据给定关于目标点的陈述,推断目标点的概率分布。如图 5(a)所示,假定参照地物为 A ,当目标对象 B 在 T 处时,观察者作出一个论断 R ,如“ B 在 A 的附近”,的概率为 P_T 。令 X 代表事件“作出论断 R ”, E_T 代表事件 B 在 T 处, P_T 可以表示为条件概率 $P(X|E_T)$ 。假定 D 为不确定性场的定义域,则对于 D 中任意一点,该条件概率都可以认为是已知的:若采用假设一,根据文献[37]的观点,该概率值表达了“附近”这个定性度量关系的模糊性,因此,可以通过合适的空间认知实验得到;而采用假设二,该概率值则可以通过隶属度的比较得到(参见图 2(b))。因此,空间定位问题就转化为计算条件概率 $P(E_T|X)$,可以采用贝叶斯定理求解该问题。

不失一般性,假定空间陈述为一个二元空间关系的简单论断,其不确定性场 $p(x, y)$ 的支持集拓扑维度为 2。将其定义域 D 划分成为 n 个等大的微分区域,其中,每个单元用 $D_i (i=1, \dots, n)$ 表示(图 5b)。因为微分单元都足够小,可以认为 D_i 是均质的,即对于 D_i 中任意点,作出论断 R 的条件概率都相等,表示为 $P(X|E_{D_i})$ 。根据贝叶斯定理,有:

$$P(E_{D_i} | X) = \frac{P(E_{D_i})P(X | E_{D_i})}{\sum_{k=1}^n P(E_{D_k})P(X | E_{D_k})} \quad (4)$$

其中, E_{D_i} 表示事件“目标点 B 位于 D_i 内部”, $P(E_{D_i} | X)$ 为 R 发生时, B 点位于 D_i 内部的条件概率, 即空间定位问题所要求解的概率。 $P(E_{D_i})$ 为关于 B 分布的先验概率, 在大多数情况下, $P(E_{D_i}) = 1/n$ 。由于对于给定论断 R 、定义域 D 和空间划分, $\sum_{k=1}^n P(X | E_{D_k})$ 为一常数, 因此, 可得 $P(E_{D_i} | X)$ 正比于 $P(X | E_{D_i})$ 。根据公式(1), 有

$$P(E_{D_i} | X) = \iint_{D_i} p(x, y) dx dy \quad (5)$$

其中, p 为描述 B 分布的概率密度函数。考虑到 D_i 足够小, 公式(5)可以转换为:

$$P(E_{D_i} | X) = p(x_0, y_0) \text{area}(D_i) \quad (6)$$

其中, (x_0, y_0) 为 D_i 中任意一点, area 为计算区域面积的函数。进而, 由于所有微分区域面积相等, 因此, 有 $p(x_0, y_0)$ 正比于 $P(X | E_{D_i})$, 而由于 D_i 的均质性, 后者等于 $P(X | E_{(x_0, y_0)})$, 其中, $E_{(x_0, y_0)}$ 表示事件“点 B 位于 (x_0, y_0) 处”。因此, 最终得到:

$$p(x_0, y_0) \propto P(X | E_{(x_0, y_0)}) \quad (7)$$

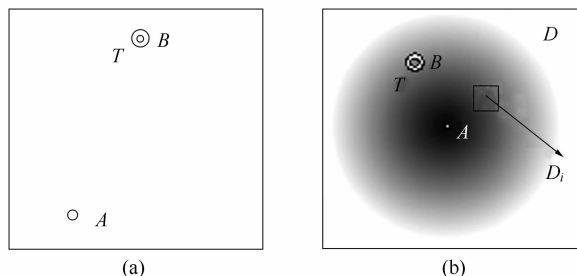


图5 不确定性场的推导示意图, 其中, A 为参照对象, B 为目标对象, D 为不确定性场的定义域, T 为 D 中任意一点, D_i 为一微分区域

Fig. 5 Demonstration of deduction of an uncertainty field, where A is the reference object, B is the target object, D is the domain of the uncertainty field, T is an arbitrary point inside D , D_i is a small differential area

以上证明结果可以概括如下, 给定一个参照对象, 对于空间中任意一点 T , 人们作出一个论断 R 的概率越高, 则当根据 R 推断目标点的位置时, 目标点位于 T 点的概率密度也越高。由于前一个概率值与相应空间关系的模糊性有关, 并可以通过设计合适的认知实验以及数学推导获得, 因此, 基于公式(7)可以计算定义域 D 中每个点 $(x_i, y_i) (i=1, \dots, n)$ 的 $p(x_i, y_i)$ 值, 从而得到空间陈述的不确定性场 $p(x, y)$ 。

4.3 应用实例

根据空间陈述对应的不确定性场模型以及文献[31]所介绍的不确定性场的操作, 本研究采用 Google Map API 开发了以自然语言陈述的空间定位工具 SA Locator (图6)。例如, 根据空间陈述“在北京大学 Geosoft 实验室以西 300m, 距离表达精度为 20m”建立其不确定性场。首先, 确定参照对象的空间位置, 根据该陈述中的地名“北京大学”和“Geosoft 实验室”, 在地名库的支持下得到地名所指称的参照对象, 而对象的地理覆盖属性记录了对象的空间位置表现为一个点坐标, 我们使用红色图钉状符号将其标识在 Google 地图中(图6)。其次, 根据该陈述中的空间关系“以西 $300 \pm 20\text{m}$ ”确定不确定性场。该空间关系可以分解为两个原子关系, 即“以西”和“距离 $300 \pm 20\text{m}$ ”。在本工具中, 方位关系采用了 8 方向锥形模型, 因此, 前者会得到一个 45° 的扇形区域。而后者则对应了一个宽度为 40m 的环形范围。因此, 叠加上述两个区域, 目标对象的可能分布呈一个 45° 圆环带。为了简单起见, 没有考虑空间关系的模糊性并假定误差是均匀分布的, 所以, 定义域 D 中任意一点的条件概率 $P(X | E_{D_i})$ 都相同。最后, 依公式(7)计算目标对象分布中每个点 $(x_i, y_i) (i=1, \dots, n)$ 的 $p(x_i, y_i)$ 值, 从而得到该空间陈述的不确定性场, 并栅格化为图6所示的区域, 即目标对象的可能分布范围。

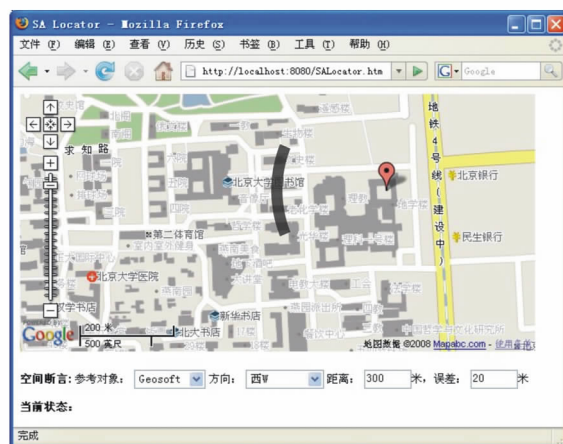


图6 基于 Google Map API 的空间定位工具 SA Locator
Fig. 6 SA Locator: a Google Map API based toolkit for positioning localities according to spatial assertions

5 结论

在地理空间知识的表达中, 地理位置主要通过

文本方式描述的。这种表达方式通常较为粗略,但是,在人们的日常交流中却发挥着重要作用。考虑到大量经济数据,以及自然观测数据中的位置也是以自然语言表达的,因此,我们需要实现对位置描述的地理定位,在本文中称之为空间定位问题。本文主要研究了空间关系和参照对象的空间陈述。在空间定位问题中,它可抽象为 3 个基本元素:目标对象、参照对象和空间关系,即目标对象和参照对象直接具有特定的空间关系。有两个假设可以用于空间定位,假设一认为目标地物使得给定陈述成立,或至少对于部分陈述者成立;假设二认为目标地物使得观察者认为给定陈述最合适。通常,地理事实和空间关系的不确定性会使得这根据两个假定得到的目标对象分布不同,而假定一会导致更大的不确定性。一个点状目标对象的分布可以用一个概率密度函数(不确定性场)表示,这使得我们可以计算目标点在二维空间中任意一个区域内的概率。最后,采用贝叶斯定理,本文证明了以下结论:对于空间中任意一点 T ,人们做出一个空间陈述的概率越高,则依该陈述,目标点位于 T 点的概率密度也越高。鉴此,我们可以根据给定的空间陈述,采用合适的空间关系模型和假设,建立相应的不确定性场表达。

致谢:郭鹏辉同学完成了文中的空间定位工具,在此表示感谢。

参考文献:

- [1] Bennett B, Agarwal P. Semantic categories underlying the meaning of "Place"[C]. // Winter S, Duckham M, Kulik L, Kuipers B (ed.). COSIT'07, Springer, 2007, 78 - 95.
- [2] Chen Z. Hierarchy of knowledge in GIS[J]. Science in China Series E-Technological Sciences, 2008, 51(S1): 6 - 12.
- [3] Zhang Y, Gao Y, Xue L, *et al.* A common sense geographic knowledge base for GIR[J]. Science in China Series E-Technological Sciences, 2008, 51(S1): 26 - 37.
- [4] Mark D M, Freksa C, Hirtle S, *et al.* Cognitive models of geographical space[J]. International Journal of Geographical Information Science, 1999, 13(8): 747 - 774.
- [5] 刘瑜, 方裕, 郭伦, 等. 基于场所的 GIS 研究[J]. 地理与地理信息科学, 2005, 21(5): 6 - 11.
- [6] Agarwal P. Ontological considerations in GIScience[J]. International Journal of Geographical Information Science, 2005, 19(5): 501 - 536.
- [7] 陈军, 赵仁亮. GIS 空间关系的基本问题与研究进展[J]. 测绘学报, 1999, 28(2): 95 - 102.
- [8] Cohn A G, Hazarika S M. Qualitative spatial representation and reasoning: An overview[J]. Fundamenta Informaticae, 2001(46): 10 - 29.
- [9] 刘瑜, 龚咏喜, 张晶, 等. 地理空间中的空间关系表达和推理[J]. 地理与地理信息科学, 2007, 23(5): 1 - 7.
- [10] Hill L L. Georeferencing—The geographic associations of information[M]. Cambridge: MIT Press, 2006.
- [11] Goodchild M F, Hill L L. Introduction to digital gazetteer research[J]. International Journal of Geographical Information Science, 2008, 22(10): 1039 - 1044.
- [12] Liu Y, Li R, Chen K, *et al.* KIDGS: A geographical knowledge-informed digital gazetteer service [C]. Geoinformatics 2009, Fairfax, 2009, 1 - 6.
- [13] Jones C B, Purves R S. Geographical information retrieval[J]. International Journal of Geographical Information Science, 2008, 22(3): 219 - 228.
- [14] MaNIS (Mammal Networked Information System), 2001, Available online at: <http://manisnet.org> (accessed 2007 - 11 - 16).
- [15] Goldberg D W, Wilson J P, Knoblock C A. From text to geographic coordinates: The current state of geocoding[J]. URISA Journal, 2007, 19(1): 33 - 46.
- [16] 刘瑜, 张毅, 田原, 等. 广义地名及其本体研究[J]. 地理与地理信息科学, 2007, 23(6): 1 - 7.
- [17] Wiecek J, Guo Q, Hijmans R J. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty[J]. International Journal of Geographical Information Science, 2004, 18(8): 745 - 767.
- [18] Guo Q, Liu Y, Wiecek J. Georeferencing locality descriptions and computing associated uncertainty in a probabilistic approach[J]. International Journal of Geographical Information Science, 2008, 22(10): 1067 - 1090.
- [19] Renz J. Qualitative spatial reasoning with topological information[M]. Berlin Heidelberg: Springer-Verlag, 2002.
- [20] 郭庆胜, 杜晓初, 刘浩. 空间拓扑关系定量描述与抽象方法研究[J]. 测绘学报, 2005, 34(2): 123 - 128.
- [21] 吴静, 程朋根, 陈斐, 等. 空间目标的方向关系定性推理[J]. 测绘学报, 2006, 35(2): 160 - 165.
- [22] Takemura C M, Cesar Jr, R M, Bloch I. Fuzzy modeling and evaluation of the spatial relation 'along'[C]. // Lazo M, Sanfeliu A (eds.). CIARP 2005, (Lecture Notes in Computer Science 3773, Berlin: Springer-

- er, 2005, 837–848.
- [23] Bloch I, Colliot O, Cesar, R M. On the ternary spatial relation “between”[J]. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 2006, 36(2):312–327.
- [24] Randell D A, Cui Z, Cohn A G. A spatial logic based on regions and connection[C]. *Proceedings of the Third International Conference of Knowledge Representation and Reasoning*, San Mateo, CA, 1992, 165–176.
- [25] Liu Y, Wang X, Jin X, *et al.* On internal cardinal direction relations[C]. // Cohn A G, Mark D M (ed.). *COSIT 2005*, Berlin Heidelberg: Springer-Verlag, 2005, 283–299.
- [26] Egenhofer M J, Mark D M. Naive geography[C]. // Frank A U, Kuhn W (ed.). *COSIT 1995*, Berlin Heidelberg: Springer-Verlag, 1995, 1–15.
- [27] Xu, J. Formalizing natural-language spatial relations between linear objects with topological and metric properties[J]. *International Journal of Geographical Information Science*, 2007, 21(4):377–395.
- [28] 金鑫, 耿海燕, 高勇, 等. 空间方位关系在不同认知场境下的模糊性探讨[J]. *北京大学学报(自然科学版)*, 2009, 45(6):1025–1032.
- [29] Gong Y, Wu L, Lin Y, *et al.* Probability issues in locality descriptions based on Voronoi neighbor relationship[J]. *Journal of Visual Languages and Computing*, 2012, 23(4):213–222.
- [30] Wang X, Liu Y, Gao Z, *et al.* Landmark-based Qualitative Reference System[C]. *Proceedings of IGARSS 2005*, Korea, Seoul, 2005(2):932–935.
- [31] Liu Y, Guo Q, Wicczorek J, *et al.* Positioning localities based on spatial assertions[J]. *International Journal of Geographical Information Science*, 2009, 23(11):1471–1051.
- [32] 刘瑜, 袁一泓, 张毅. 基于认知的模糊地理要素建模——以中关村为例[J]. *遥感学报*, 2008, 12(2):370–377.
- [33] Dilo A, De By, R A, *et al.* A system of types and operators for handling vague spatial objects[J]. *International Journal of Geographical Information Science*, 2007, 21(4):397–426.
- [34] Worboys M F, Clementini E. Integration of imperfect spatial information[J]. *Journal of Visual Languages and Computing*, 2001(12):61–80.
- [35] Dempster A P. A generalization of Bayesian inference [J]. *Journal of the Royal Statistical Society, Series B*, 1968(30):205–247.
- [36] Shafer G. *A Mathematical Theory of Evidence*[M]. Princeton, NJ: Princeton University Press, 1976.
- [37] Dubois T, Prade H. Fuzzy sets and probability: Misunderstandings, bridges and gaps[C]. *Proceedings of 2nd IEEE Conf. on Fuzzy Systems*, 1993, 1059–1068.

On the Description-based Spatial Positioning and the Associated Uncertainty

ZHANG Yi¹, WU Yang², GAO Yong¹ and LIU Yu¹

(1. *Institute of Remote Sensing and Geographical Information Systems, Peking University, Beijing 100871, China*;
2. *School of Geomatics and Urban Information, Beijing University of Civil Engineering and Architecture, Beijing 100044, China*)

Abstract: In the representation and communication of geographical knowledge of human beings, locations are often textually expressed. In addition to place names and addresses, a spatial assertion, which consists of reference objects and spatial relations, can roughly describe a locality. Typical examples of spatial assertions are “the vicinity of Peking University” and “Between Peking University and Tsinghua University”. The objective of spatial positioning is to infer the location of a target object and the associated uncertainty for a given spatial assertion. In this research, we investigate four levels of uncertainty associated with spatial assertions. They are uncertainty of spatial assertions, imperfection of reference objects, imprecision and vagueness of spatial relations, and distribution of target objects. If a target object is abstracted to a point, the concept of an uncertainty field can be used to represent the possible spatial distri-

bution of the target object. An uncertainty field is a 2-dimensional probability density function that satisfies $\iint_F p(x,y)dx dy = 1$ where $p(x,y)$ denotes the field and F is its support. For a point T and a reference object A , if the probability that an observer make a description X is high, then the probability density at O is also high when X is known. This argument is proved using Bayes' theorem and can be viewed as the theoretical foundation to establish an uncertainty field. We also implemented a toolkit to generate uncertainty fields based on a number of spatial assertions.

Key words: spatial assertion; spatial positioning; reference object; target object; uncertainty