DOI:10.13203/j.whugis2004.11.019
第29卷 第11期 武汉大学学报·信息科学版 Vol.29 No.11
2004年11月 Geomatics and Information Science of Wuhan University Nov. 2004

文章编号:1671-8860(2004)11-1021-04 文献标识码:A

宽边界区域拓扑关系量化分析与抽象

杜晓初[1] 郭庆胜[1] 丁虹[1]

(1 武汉大学资源与环境科学学院,武汉市珞喻路129号,430079)

摘 要:对Cohn等人提出的卵-黄模型所确定的宽边界区域的46种拓扑关系及其聚类方法进行了分析,提出了一种定量化方法对这些拓扑关系进行描述,并对这46种拓扑关系进行了抽象。

关键词:宽边界区域;空间拓扑关系;定量化分析
中图法分类号:P283.1

对于宽边界区域的拓扑关系,很多学者进行了研究。Cohn等人基于区域对模型建立了一种描述不确定性区域的空间拓扑关系模型,称为卵-黄(egg-yolk)模型[1]。卵黄相当于区域的内部,卵白相当于区域的宽边界。这种模型利用两个区域的卵和卵黄之间的RCC-5关系,对区域之间的拓扑关系进行分类,可以确定46种拓扑关系。对宽边界区域的这种分类当然是十分必要的,但在本质上这种分类还比较模糊,应对其进行进一步的分类。利用定量化分析方法是解决这类问题的最佳途径。

1 宽边界区域拓扑关系的定量化

对宽边界区域拓扑关系的定量化,已经有很多学者进行了研究。有些学者在考察确定的模糊拓扑关系中引入了9交集合的度量[2];Schneider利用离散模糊区域的隶属函数,可以方便地计算两个模糊区域之间不同拓扑关系的隶属度[3]。

笔者采用类似的方法,用定量化的空间关系向量模型来描述宽边界区域之间的拓扑关系,该模型的原型是拓扑关系的4-交集模型。

$$R_M=(m(\tilde{A}^0\cap B^0),m(\tilde{A}^0\cap\partial\tilde{B}),m(\partial\tilde{A}\cap B^0),m(\partial\tilde{A}\cap\partial\tilde{B}))\quad(1)$$

式中,$m(\tilde{A}^0\cap B^0)=\dfrac{\mathrm{Area}(\tilde{A}^0\cap B^0)}{\min\{\mathrm{Area}(\tilde{A}^0),\mathrm{Area}(B^0)\}}$

$$m(\tilde{A}^0\cap\partial\tilde{B})=\frac{\mathrm{Area}(\tilde{A}^0\cap\partial\tilde{B})}{\min\{\mathrm{Area}(\tilde{A}^0),\mathrm{Area}(\partial\tilde{B})\}}$$

$$m(\partial\tilde{A}\cap B^0)=\frac{\mathrm{Area}(\partial\tilde{A}\cap B^0)}{\min\{\mathrm{Area}(\partial\tilde{A}),\mathrm{Area}(B^0)\}}$$

$$m(\partial\tilde{A}\cap\partial\tilde{B})=\frac{\mathrm{Area}(\partial\tilde{A}\cap\partial\tilde{B})}{\min\{\mathrm{Area}(\partial\tilde{A}),\mathrm{Area}(\partial\tilde{B})\}}$$

其中,$m(\ )$为两个宽边界区域$\tilde{A}$、B的内部和宽边界$\tilde{A}^0$、B^0、$\partial\tilde{A}$和$\partial\tilde{B}$分别相交的程度;Area( )为求区域面积的运算符;min{}为取两者中较小的面积。由此定义可以知道,这种定量化关系向量中的每一个分量的值都在[0, 1]之间,可以很好地表示两个宽边界区域的内部和宽边界两两相交的程度。

2 宽边界区域拓扑关系的抽象

空间拓扑关系抽象是人类空间认知和思维的基本方法之一,抽象的结果必须更易于理解和表达以及进行信息的交流,因此必须有一个标准。

2.1 参考空间拓扑关系

Egenhofer和Franzosa等人提出的4-交集模型可以确定区域之间的8种拓扑关系:相离、相接、相交、相等、包含、被包含、覆盖和被覆盖[4]。这8种关系与RCC-8关系集中的8种关系是完全一致的,因此,也可以用来代替RCC-5关系集中的5种关系。但是考虑到这8种拓扑关系中的覆盖与被覆盖、包含与被包含两对关系是对称的,只是参考对象不同,因此只考虑包含和覆盖关系即可。这样,上面的8种拓扑关系可以简化为6种:相离、相接、相交、相等、包含和覆盖。

收稿日期:2004-06-29。
项目来源:国家教育部高校骨干教师基金资助项目(2000G07)。

为了方便空间拓扑关系的抽象,这里用一个4维空间向量来描述4-交集模型:

$$R=(A^0\cap B^0, A^0\cap \partial B, \partial A\cap B^0, \partial A\cap \partial B) \quad (2)$$

式中,A^0和B^0、∂A和∂B分别表示确定集合A、B的内部和边界。

这样,6种参考空间拓扑关系的向量形式分别为:相离:$R_d=(0,0,0,0)$;相接:$R_m=(0,0,0,1)$;相交:$R_o=(1,1,1,1)$;相等:$R_e=(1,0,0,1)$;覆盖:$R_c=(1,1,0,1)$;包含:$R_{ct}=(1,1,0,0)$。笔者称这6个向量为参考空间关系向量。

2.2 空间向量的相关度

定义 设S、T是两个n维空间向量,称

$$r(S,T)=1-\frac{1}{n}\sum_{i=1}^{n}\|x_{si}-x_{ti}\| \quad (3)$$

为空间向量S、T的相关度。式中,x_{si}和x_{ti}分别表示空间向量S和T的第i个分量,且$0\leqslant x_{si}, x_{ti}\leqslant 1$;$\|*\|$表示这两个向量的两个分量之间的欧氏距离,则两个分量之间的最大距离为1。

根据空间向量相关度的定义可以知道,相关度越大,两个向量相关性越大,也就是说,这两个向量更为类似。这样,分别计算一个被考察的空间向量与几种作为标准的参考空间向量的相关度,然后比较计算结果,就可以将这个空间向量进行归类。在本文中,把计算两个空间向量的相关度作为宽边界区域拓扑关系抽象的基本方法。

2.3 宽边界区域拓扑关系抽象的基本方法

宽边界区域拓扑关系的抽象可以按照以下的几个步骤进行。

1) 计算定量化空间关系向量的值。根据公式(1),对于考察的一对宽边界区域,可以计算出这一对区域的定量化空间关系向量的值R_M;

2) 计算相关度r。根据式(3),分别计算定量化空间关系向量的值R_M与6个参考关系向量R_d、R_m、R_o、R_e、R_c和R_{ct}的相关度$r(R_M, R_d)$、$r(R_M, R_m)$、$r(R_M, R_o)$、$r(R_M, R_e)$、$r(R_M, R_c)$和$r(R_M, R_{ct})$。

3) 判定宽边界区域的拓扑关系。比较6个相关度的大小,与R_M相关度最大的那一种拓扑关系即为被考察的这一对宽边界区域之间的拓扑关系。例如,如果算出$r(R_M, R_o)$的值最大,则认为这一对宽边界区域的拓扑关系是相交的。

这样,就可以将宽边界区域之间的46种拓扑关系抽象为基本的6种拓扑关系。如果考虑到这6种拓扑关系所属的定量化空间关系向量的分布特征,就可以在前面抽象的基础上进行进一步的抽象,以增强抽象的准确性。

2.4 宽边界区域拓扑关系的再聚类

前面描述的宽边界区域拓扑关系抽象的基本方法是以6种基本的空间关系向量作为参考,并没有考虑到这些拓扑关系定量化计算以后的数量分布特征可能导致关系判定的误差。如果结合这46种拓扑关系定量化以后的数量分布特征进行抽象,就可以避免这种误差。这种再聚类的步骤如下。

1) 计算每一对宽边界区域之间的定量化关系向量值。在这里综合考虑这46种拓扑关系,因此,首先计算这46种拓扑关系的定量化空间关系向量:

$$R_M^i=(m_i(A^0\cap B^0), m_i(A^0\cap \partial B), m_i(\partial A\cap B^0), m_i(\partial A\cap \partial B))\ (1\leqslant i\leqslant 46) \quad (4)$$

2) 计算相关度r,并判定每一对宽边界区域的拓扑关系。分别求出这46个定量化空间关系向量$R_M^i(1\leqslant i\leqslant 46)$与6个参考空间关系向量$R_d$、$R_m$、$R_o$、$R_e$、$R_c$、$R_{ct}$的相关度$r(R_M^i, R)$,通过比较其相关度的大小,分别将这46种拓扑关系进行归类。这样,可以将宽边界区域的46种拓扑关系抽象为6种基本关系。

3) 重新确定参考空间关系向量。设每一种基本关系中包含了宽边界区域的拓扑关系k_j种$(1\leqslant j\leqslant 6)$,显然,$k_1+k_2+\cdots+k_6=46$。现在重新确定这6种基本空间拓扑关系的参考空间关系向量。根据聚类的思想,认为这6种基本空间拓扑关系是包含一定数量定量化空间关系向量的类,这样,就可以这6个类的中心作为新的参考空间关系向量。根据中心计算公式:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

可得到新的参考空间关系向量(式中,n为被考察元素的个数;x_i为被考察元素的坐标):

$$R_M(j)=(\overline{m}_j(A^0\cap B^0), \overline{m}_j(A^0\cap \partial B), \overline{m}_j(\partial A\cap B^0), \overline{m}_j(\partial A\cap \partial B))\ (1\leqslant j\leqslant 6) \quad (5)$$

这里,向量的每一个分量分别定义为:

$$\overline{m}_j(A^0\cap B^0)=\frac{1}{k_j}\sum_{i=1}^{k_j}m_i(A^0\cap B^0),$$

$$\overline{m}_j(A^0\cap \partial B)=\frac{1}{k_j}\sum_{i=1}^{k_j}m_i(A^0\cap \partial B),$$

$$\overline{m}_j(\partial A\cap B^0)=\frac{1}{k_j}\sum_{i=1}^{k_j}m_i(\partial A\cap B^0),$$

$$\overline{m}_j(\partial A\cap \partial B)=\frac{1}{k_j}\sum_{i=1}^{k_j}(\partial A\cap \partial B),$$

这样就可得到 6 类基本拓扑关系的新参考空间关系向量 $R_M(1)$, $R_M(2)$, …, $R_M(6)$, 分别对应 R_d、R_m、R_o、R_e、R_c、R_{ct}。从上面的定义可以看出,新的参考空间关系向量事实上就是这已经分出的 6 种基本拓扑关系中所包含的定量化空间关系向量的平均值。

4) 拓扑关系的再聚类。用这些新的参考空间关系向量来分别代替原来的参考空间关系向量 R_d、R_m、R_o、R_e、R_c、R_{ct},分别求出每一个定量化空间关系向量 R_M^i($1 \leqslant i \leqslant 46$)与这 6 个新的参考空间关系向量 $R_M(1)$、$R_M(2)$、…、$R_M(6)$的相关度,根据相关度的大小重新将这 46 种宽边界拓扑关系归类,完成拓扑关系的再聚类。

3　应用举例

下面通过一个实例来说明宽边界拓扑关系抽象方法。在这 46 种关系中各选取一个样本,分别计算其度量化空间关系向量和相关度,其计算结果如表 1 所示。

表 **1**　试验中选定的 **46** 种拓扑关系的相关度计算表

Tab. 1　Correlation Degrees Between the 46 Topological Relations and the 6 Reference Spatial Relations

关系	定量化空间关系向量	相关度						关系	定量化空间关系向量	相关度					
		R_d	R_m	R_o	R_e	R_c	R_{ct}			R_d	R_m	R_o	R_e	R_c	R_{ct}
1	(0.00, 0.00, 0.00, 0.00)	$\underline{1.00}$	0.75	0.00	0.50	0.25	0.50	24	(1.00, 1.00, 0.00, 0.00)	0.50	0.25	0.50	0.50	0.75	$\underline{1.00}$
2	(0.00, 0.00, 0.00, 0.35)	$\underline{0.91}$	0.84	0.09	0.59	0.34	0.41	25	(1.00, 0.55, 0.00, 0.65)	0.45	0.52	0.55	0.78	$\underline{0.80}$	0.72
3	(0.00, 0.00, 0.38, 0.40)	$\underline{0.80}$	0.75	0.20	0.50	0.35	0.40	26	(1.00, 0.40, 0.00, 0.80)	0.45	0.60	0.55	$\underline{0.85}$	0.80	0.65
4	(0.00, 0.48, 0.00, 0.60)	0.73	$\underline{0.88}$	0.27	0.53	0.52	0.47	27	(1.00, 0.70, 0.00, 0.40)	0.48	0.42	0.52	0.68	0.78	$\underline{0.82}$
5	(0.00, 0.00, 1.00, 0.56)	0.61	$\underline{0.64}$	0.39	0.39	0.11	0.11	28	(0.00, 1.00, 1.00, 0.80)	0.30	0.45	$\underline{0.70}$	0.20	0.45	0.30
6	(0.00, 0.00, 1.00, 0.44)	$\underline{0.64}$	0.61	0.36	0.36	0.14	0.14	29	(0.65, 0.45, 0.00, 0.75)	0.55	0.64	0.45	$\underline{0.81}$	0.69	0.59
7	(0.00, 0.00, 1.00, 1.00)	0.50	$\underline{0.75}$	0.50	0.50	0.25	0.00	30	(1.00, 0.75, 0.00, 0.00)	0.56	0.31	0.44	0.56	0.69	$\underline{0.94}$
8	(000, 0.00, 1.00, 1.00)	0.50	$\underline{0.75}$	0.50	0.50	0.25	0.00	31	(1.00, 0.78, 0.00, 0.40)	0.46	0.41	0.54	0.65	0.80	$\underline{0.84}$
9	(0.00, 0.44, 0.56, 0.44)	$\underline{0.63}$	0.62	0.37	0.37	0.34	0.35	32	(1.00, 0.72, 0.00, 0.46)	0.45	0.47	0.55	0.68	0.79	$\underline{0.81}$
10	(0.00, 0.35, 1.00, 0.45)	$\underline{0.56}$	0.51	0.44	0.26	0.19	0.24	33	(1.00, 0.60, 0.00, 0.40)	0.50	0.45	0.50	0.70	0.75	$\underline{0.80}$
11	(0.00, 0.60, 1.00, 0.60)	0.45	0.50	$\underline{0.55}$	0.25	0.30	0.25	34	(0.00, 1.00, 1.00, 0.74)	0.31	0.43	$\underline{0.69}$	0.43	0.43	0.31
12	(0.00, 0.40, 1.00, 0.90)	0.44	$\underline{0.63}$	0.56	0.38	0.30	0.13	35	(0.44, 0.56, 0.52, 0.76)	0.43	0.56	$\underline{0.57}$	0.53	0.56	0.43
13	(0.00, 0.30, 1.00, 0.90)	0.44	$\underline{0.65}$	0.56	0.40	0.40	0.10	36	(1.00, 0.65, 0.00, 1.00)	0.34	0.59	0.66	0.86	$\underline{0.91}$	0.66
14	(0.40, 0.40, 0.50, 0.30)	$\underline{0.60}$	0.50	0.40	0.45	0.40	0.50	37	(1.00, 1.00, 0.00, 0.00)	0.50	0.25	0.50	0.50	0.75	$\underline{1.00}$
15	(0.60, 0.60, 0.80, 0.56)	0.36	0.44	$\underline{0.64}$	0.44	0.49	0.46	38	(1.00, 0.72, 0.00, 1.00)	0.32	0.57	0.68	0.82	$\underline{0.93}$	0.68
16	(0.60, 0.80, 0.65, 0.58)	0.34	0.38	$\underline{0.66}$	0.43	0.58	0.54	39	(1.00, 0.00, 0.00, 0.54)	0.61	0.63	0.39	$\underline{0.88}$	0.63	0.61
17	(1.00, 0.61, 0.00, 0.62)	0.45	0.50	0.55	0.55	$\underline{0.81}$	0.25	40	(1.00, 0.00, 0.00, 1.00)	0.50	0.75	0.50	$\underline{1.00}$	0.75	0.50
18	(1.00, 0.69, 0.00, 0.58)	0.43	0.47	0.57	0.72	$\underline{0.82}$	0.78	41	(1.00, 0.00, 0.00, 1.00)	0.50	0.75	0.50	$\underline{1.00}$	0.75	0.50
19	(0.00, 1.00, 1.00, 0.60)	0.35	0.40	$\underline{0.65}$	0.40	0.40	0.35	42	(0.00, 1.00, 1.00, 1.00)	0.25	0.50	$\underline{0.75}$	0.25	0.50	0.50
20	(0.40, 0.60, 0.70, 0.65)	0.41	0.49	$\underline{0.59}$	0.49	0.48	0.33	43	(0.80, 0.60, 0.70, 0.90)	0.25	0.45	$\underline{0.75}$	0.60	0.60	0.45
21	(0.75, 0.51, 0.70, 0.80)	0.32	0.47	$\underline{0.68}$	0.59	0.58	0.33	44	(1.00, 0.60, 0.00, 1.00)	0.35	0.60	0.65	0.85	$\underline{0.90}$	0.65
22	(0.50, 0.55, 0.45, 0.72)	0.45	0.53	0.55	0.53	$\underline{0.58}$	0.47	45	(1.00, 0.65, 0.00, 1.00)	0.34	0.59	0.66	0.84	$\underline{0.91}$	0.66
23	(1.00, 1.00, 0.00, 0.00)	0.50	0.25	0.50	0.50	0.75	$\underline{1.00}$	46	(1.00, 0.00, 0.00, 1.00)	0.50	0.75	0.50	$\underline{1.00}$	0.75	0.50

注:表中有下划线的数为最大相关度的值。

根据计算结果,把归类结果与前文提到的 Cohn 的 13 种归类进行比较,发现仅仅当两个宽边界区域各部分相交较少时,与"相交"这一类不完全一致,但是当相交程度较多时是一致的;与 Clementini 等人的 14 种归类相比较,也只有与其中的 meet 和 overlap 不完全符合。分析其原因,这两种归类方法都以两个宽边界区域的各部分是否相接或相交作为判定的依据,而没有考虑到相接或相交程度的大小。

通过计算重叠区域的相交程度来实现拓扑关系的抽象,存在从量变到质变的过程。如果以 6 个参考空间关系向量作为参考,发现 0.5 是产生拓扑关系变化的一个分界线,也就是说,当两个定量化空间关系向量的一个分量的值大于和小于 0.5 时,如果其余的值大体保持不变,那么这两种关系可能是概念邻近图中相邻的两种拓扑关系。如表 1 中的关系 5 和 6、21 和 22、25 和 26 都是如此。

在该实验中最后得出的各个类型的拓扑关系数分别为:相离(7)、相接(6)、相交(11)、相等(6)、覆盖(8)、包含(8)。根据式(5)可以分别求出新的参考空间关系向量:

$$R_M(1) = (0.06,\ 0.17,\ 0.35,\ 0.33);$$

$$R_M(2) = (0.00,\ 0.15,\ 0.83,\ 0.79);$$

$$R_M(3) = (0.33,\ 0.75,\ 0.82,\ 0.72);$$

$$R_M(4) = (0.94,\ 0.14,\ .00,\ 0.85);$$

$$R_M(5) = (0.94,\ 0.72,\ 0.06,\ 0.82);$$

$R_M(6) = (1.00, 0.82, 0.00, 0.21)$

很明显,这些新的参考空间关系向量与原来的6个参考空间关系向量有很强的相关性,可以以这些向量作为参考对这46种拓扑关系进行重新抽象,以提高准确性。

空间拓扑关系的这种抽象在实际中经常见到,如判定一所房子和一片(逐渐扩展的)林地的拓扑关系;判定两个城市的影响区域的关系等。

4 结 语

本文提出一种定量化的方法对宽边界区域的46种拓扑关系进行抽象。这种方法的主要思想是用两个宽边界区域的内部和宽边界以及它们两两相交的交集的面积,来表征这些交集的相交程度;然后通过计算这些定量化的空间关系向量与参考空间关系向量的相关程度,来确定拓扑关系的类别。运用这种方法,可以将这46种拓扑关系抽象为最基本的6种拓扑关系。类似的定量化分析方法也可以用于方向关系、相似关系等空间关系的研究,这也是今后研究的方向。

参 考 文 献

1 Cohn A G, Gotts N M. The "Egg-yolk" Representation of Regions with Indeterminate Boundaries. In: Burrough P A, Frank A U, eds. Geographic Objects with Indeterminate Boundaries (GISDATA 2). London: Taylor & Francis, 1996. 172~187

2 Cheng X Y, Takeshi, Nasu M. Spatial Relations Between Uncertain Sets. International Archives of Photogrammetry and Remote Sensing, 1996, 31(B3): 105~110

3 Schneider M. Finite Resolution Crisp and Fuzzy Spatial Objects. The 9th International Symposium on Spatial Data Handing, Beijing, 2000

4 Egenhofer M, Franzosa R. Point-Set Topological Spatial Relations. International Journal of Geographical Information Systems, 1991, 5(2): 161~174

5 郭庆胜, 郑春燕. 地图线状符号图案单元的优化配置方法. 武汉大学学报·信息科学版, 2002, 27(5): 499~504

第一作者简介:杜晓初,博士生。主要研究方向为地理信息不确定性分析、地理空间推理及其应用。
E-mail: duxiaochu@sina.com

Quantitative Analysis and Abstraction of Topological Relation Between Regions with Broad Boundaries

DU Xiaochu[1] *GUO Qingsheng*[1] *DING Hong*[1]

(1 School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: Topological relations between regions with broad boundaries are very complicated. It is difficult to distinguish these topological relations by qualitative method. This paper analyzes the 46 topological relations between regions with broad boundaries determined by the "egg-yolk" model proposed by Cohn and others and the clustering of these relations and proposes a quantitative method that can describe these topological relations.

Key word: region with broad boundary; spatial topological relations; quantitative analysis

About the first author: DU Xiaochu, Ph. D candidate, majors in analysis of uncertainty for geographical information, spatial reasoning and its application.
E-mail: duxiaochu@sina.com

(责任编辑: 洪光)