

空间分析组件的研究与设计

杨占龙 汪林林

(重庆邮电学院 重庆400065)

Research and Design about Component of Spatial-Analyzer

YANG Zhan-Long WANG Lin-Lin

(Chongqing University of Posts & Telecommunication, Chongqing 400065)

Abstract Components GIS(ComGIS), a new technology based on Components Software, is a mainstream in GIS technology. In this paper, the phases in which software of GIS develops are discussed in detail, and the merits and trends of ComGIS are pointed out. Also, expatiates how to design the component of Spatial-Analyzer function in ComGIS.

Keywords Geographic information system, ComGIS, Component, Spatial-analyzer

1. GIS 的组件化趋势

地理信息系统(GIS)技术正处于一个重要的发展时期,新概念和新产品层出不穷。在GIS蓬勃发展的今天,GIS的组件化趋势日益明显,已经成为GIS的重要发展方向之一。从发展历程看,GIS可以划分为图1所示的几个发展阶段。

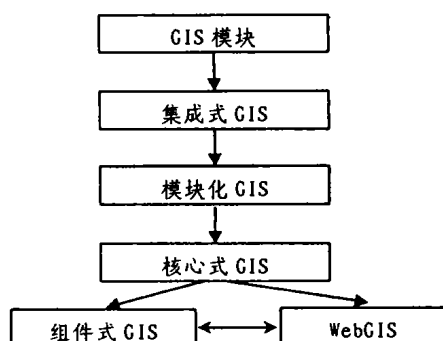


图1 GIS软件发展历程

在GIS发展的早期阶段,由于受到技术的限制,GIS软件

往往是只能满足于某些功能要求的一些模块,没有形成完整的系统,而各个模块之间不具备协同工作的能力。

随着理论和技术的发展,各种GIS模块走向集成,逐步形成大型GIS软件包(GIS Package),我们暂称之为集成式GIS(IntegratedGIS),如ESRI的Arc/Info、Genasys的GenaMap等均为集成式GIS的代表。

另一类GIS为随后出现的模块化GIS(Modular GIS),代表软件有Intergraph的MGE等。模块化GIS的基本思想是把GIS按照功能划分为一系列模块,运行于统一的基础环境之上(如MicroStation)。但无论是集成式GIS或是模块化GIS,都很难与管理信息系统(MIS)以及专业应用模型集成高效、无缝的GIS应用。

为解决集成式GIS与模块化GIS的缺点,提出了核心式GIS(Core GIS)的概念。但核心式GIS被设计为操作系统的基本扩展,也不适应可视化程序设计的潮流。

随着计算机软件技术的发展,GIS组件化发展到了一个全新的阶段,出现了组件式GIS(Components GIS,缩写为ComGIS)。组件式GIS基于标准的组件式平台,各个组件之间不仅可以进行自由、灵活的重组,而且具有可视化的界面和

杨占龙 硕士研究生,从事GIS系统、数据库系统、计算机算法分析、计算机网络等方面的研究。汪林林 教授,从事GIS系统、数据库系统、计算机算法分析、计算机网络等方面的研究。

我们的核酸序列数据库上联机分析处理的界面如图3所示。

结论 本文针对当前广泛采用的EMBL核酸序列数据库,提出了有针对性的联机分析处理实现方案。该方案给出了从文件格式的数据到关系数据库再到OLAP多维数据集的一系列数据转存实现方法。并利用转存之后的数据进行了核酸序列分布规律和各类别核酸序列组成成份的统计分析。该方案可以有效地解决原始数据文件分析速度慢,分析能力差的弱点,可以被视为将联机分析处理(OLAP)技术应用于生物信息学的一次成功探索。另外,该方案也完全适用于蛋白质数据库(如Swiss-Prot)。

参考文献

1 Inmon W H. Building the Data Warehouse Second Edition. John

Wiley&Sons Inc, 1996

2 Melinda G, Song, Il-Yeol. Data warehouse design for pharmaceutical drug discovery research Axel. In: Intl. Conf. on Database and Expert Systems Applications - DEXA. Sep. 1997

3 Xie Guochun, DeMarco R, Blevins R, Wang Yuhong. Storing biological sequence databases in relational form Bioinformatics 16: 288~289

4 Microsoft Corporation Microsoft SQL Server 2000 Analysis Services Microsoft Press, 2000

5 Stoesser G. The EMBL nucleotide sequence database. Nucleic Acids Research, 2001, 29(1)

6 European Bioinformatics Institute. EMBL Nucleotide Sequence Database: User Manual Release 70, March 2002

使用方便的标准接口。组件式 GIS 代表着当今 GIS 发展的潮流。

几乎在开发组件式 GIS 的同时,出现了万维网 GIS(WebGIS)。万维网 GIS 是 Internet 技术与 GIS 相结合的产物。从基础软件的角度来看,组件式 GIS 和万维网 GIS 之间的界限将会变得模糊不清,而只是应用方式上的区别。

从 GIS 模块发展到集成式 GIS 是从分散到集中的过程,这是 GIS 发展历程中的一个重大进步。从集成式 GIS 发展到模块化 GIS,这是 GIS 组件化的开始,随后发展到核心式 GIS, GIS 组件化趋势越来越明显,并形成组件化的标准形式——组件式 GIS 和 WebGIS。

2. COM 的优势

COM(Component Object Model)是 Windows 操作系统的技术和核心,是搭建 Windows 应用程序的基础。COM 提供了二进制级别的代码复用机制。它提供了一种基于查询的机制来完成对象之间的通讯,以客户/服务器程序的体系结构协同完成特定的功能。COM 提供了以下的特性:

- 同语言的无关性:二进制的设计标准使得基于 Windows 平台的开发工具可以混用 COM 组件,比如你可以用 C++ 编写组件,而在 VB 中使用该组件。

- 软件版本升级的健壮性:COM 组件通过提供多接口机制支持软件版本升级的健壮性,COM 组件的不同功能用不同的接口向外展示。要为旧的组件添加新的功能时,为组件新加一个接口就行了。在这种情况下,允许旧的应用程序在不进行更改的情况下运行,同时新的应用程序可以通过查询新接口利用组件的新功能。

- 面向对象的特性:COM 允许软件模块以面向对象的方式传递其功能。对于 C++ 程序员而言,面向对象的特征是最为常见的。COM 提供三种基本的面向对象的特征,即封装、继承和多态,并且 COM 是以一种与语言无关的方式对这三种特征提供了支持。

- 位置透明性:该特性意味着组件的用户并不需要明确地了解组件所处的位置,这个组件可能在一个 DLL 中(进程内服务)、可能在一个 EXE 中(进程外服务)、也可能在一个远程计算机上运行(分布式组件对象)。但这些位置对于客户程序来说是透明的,客户程序在使用组件对象模型时不需要考虑该组件所处的位置。

3. 空间分析组件的设计

空间分析早已成为地理信息系统的核心功能之一,它特有的对地理信息(特别是隐含信息)的提取、表现和传输功能,是地理信息系统区别于一般信息系统的主要功能特征。空间分析是对分析空间数据有关技术的统称。根据作用的数据性质不同,可以分为:①基于空间图形数据的分析运算;②基于非空间属性数据的运算;③空间和非空间数据的联合运算。空间分析赖以进行的基础是地理空间数据库,其运用的手段包括各种几何的逻辑运算、数理统计分析、代数运算等数学手段,最终的目的是解决人们所涉及到地理空间的实际问题,提取和传输地理空间信息,特别是隐含信息,以辅助决策。空间分析包含丰富的功能,包括几何测量、缓冲区分析、地图分析、地形分析、网络分析以及多边形操作等等。为了提高整个系统的灵活性和可靠性,我们把空间分析组件划分成多个功能组件,然后再用这些功能组件来搭建空间分析组件,以下就是空

间分析组件的功能图(图2)。

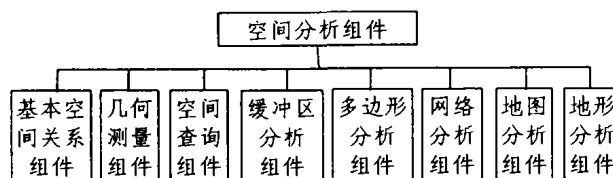


图2 空间分析组件的功能模块划分

1. 基本空间关系组件 它是整个空间分析组件的基础,它的主要用途是给空间分析组件内部其它组件提供算法支持。它提供了空间关系中最基本的八种关系运算符,完成对任意两个空间对象的空间关系的分析。这些空间关系主要有:相等(Equals),相离(Disjoint),相交(Intersects),相邻(Touches),交叉(Crosses),在内部(In),包含(Contains)和重叠(Overlaps)。所以该组件向外提供一个接口:ISRelation(基本空间分析接口)和以上八个方法(即函数),通过该接口同外界交换信息,采用查询的方式完成对这八个函数中某个函数的调用,以判断出空间对象的基本空间关系。以下就是该组件中各个方法(即函数)的描述:

- TestDisjoint()函数实现对两个空间对象的相离关系的判断,如果两个空间对象相离则返回真,否则返回假

- TestContain()函数实现对两个空间对象的包含关系的判断,如果第一个空间对象包含第二个空间对象则返回真,否则返回假

- TestEqual()函数实现对两个空间对象的相等关系的判断,如果两个空间对象相等则返回真,否则返回假

- TestOverlap()函数实现对两个空间对象的覆盖关系的判断,如果两个空间对象覆盖则返回真,否则返回假

- TestTouch()函数实现对两个空间对象的相邻关系的判断,如果两个空间对象相邻则返回真,否则返回假

- TestCross()函数实现对两个空间对象的交叉关系的判断,如果两个空间对象交叉则返回真,否则返回假

- TestIntersect()函数实现对两个空间对象相交关系的判断,如果两个空间对象相交则返回真,否则返回假

- TestIn()函数实现对一个空间对象在另一个空间对象之中的关系的判断,如果第一个空间对象在第二个空间对象中的话返回真,否则返回假

TestDisjoint()和 TestIntersect()这两个函数是完全相对的,即把其中的一个函数取反就得到了另一个函数,而 TestContain()和 TestIn()两个函数也是可逆的,即第一个空间对象在第二个空间对象中的话,那么第二个空间对象必然包含第一个空间对象。

2. 几何测量组件 几何测量是空间分析中最基本的操作之一,它包括距离的测量、面积的测量以及频率的测量等等。所以该组件向外界提供一个接口:IGeoMeasure(几何测量接口),它负责 Distance-Line() (测量直线距离), Distance-Polyline() (测量弧段距离), Area-Polygon() (测量多边形面积)等内部函数同外界交换信息。以下就是各个方法的描述:

- Distance-Line()函数实现对两点之间距离的测量,输入两个点作为参数,返回这两点之间的距离

- Distance-Polyline()函数实现对一条弧线距离的测量,输入一条弧线作为参数,返回这条弧线的总距离

- Area-Polygon()函数实现对任意多边形的面积的测

量,输入一个多边形,返回该多边形的面积。

3. 空间查询组件 空间查询同几何测量一样,都是 GIS 中最常用的功能,是地理信息系统进行高层次分析的基础。空间查询的内容很多,包括几何参数查询,如两点间的距离查询,一个多边形的周长等查询;空间定位查询,它是给定一个点或一个几何图形,检索出该图形范围内的空间对象以及相应的属性。它又包括(1)按点的查询:给定一个鼠标点,检索出离它最近的空间对象,并显示它的属性。(2)按矩形查询:给定一个矩形窗口,查询出窗口内某一类地物的所有对象。(3)按圆查询:给定一个圆或椭圆,检索出该圆或椭圆范围内的某个类或某一层的空间对象。(4)按多边形查询:用鼠标给定一个多边形,或者在图上选定一个多边形对象,检索出位于该多边形内的某一类或某一层的空间地物;空间关系查询,它包括空间拓扑关系查询和缓冲区查询。它又分为(1)邻接查询,如查询与面状地物 A 相邻的所有多边形。(2)包含关系查询,查询某一面状地物所包含的某一类的空间对象。(3)穿越查询,如查询某条公路或河流穿越了哪些省,哪些县。(4)落入查询,例如查询某个煤矿落入了哪个省。(5)缓冲区查询,它同缓冲区分析不一样,它不对原有图形进行切割,只是根据用户需要给定一个点缓冲,线缓冲或面缓冲的距离,从而形成一个缓冲区的多边形,再根据前面所述的多边形检索的原理,检索出该缓冲区内空间地物。所以该组件向外界提供三个接口:IParallelSearch(参数查询接口)、IPositionSearch(空间定位查询接口)和 IRelationSearch(空间关系查询接口),分别实现各种查询接口所包含的查询方法(即函数):PolygonPerimeter() (多边形周长)、PointSearch() (按点查询)、RectangleSearch() (按矩形查询)、RoundSearch() (按圆查询)、PolygonSearch() (按多边形查询)、AdjacentSearch() (邻接查询)、ContainSearch() (包含查询)、CrossSearch() (穿越查询)、WithinSearch() (落入查询)、BufferSearch() (缓冲查询)。以下就是该组件中各个方法的描述:

• PolygonPerimeter() 函数实现对某个多边形周长的计算,输入一个多边形,返回该多边形的周长

• PointSearch() 函数实现按点查询的功能,输入一个点,返回该点所属的空间对象或是离该点最近的空间对象

• RectangleSearch() 函数实现按矩形查询的功能,输入一个矩形窗口,返回该矩形窗口内的所有空间对象

• RoundSearch() 函数实现按圆查询的功能,输入一个圆的圆心和半径,返回该圆范围内的所有空间对象

• PolygonSearch() 函数实现按多边形的查询功能,输入一个多边形,返回该多边形范围内的所有空间对象

• AdjacentSearch() 函数实现邻接查询的功能,输入任意一个空间对象,返回与该对象邻接的所有空间对象

• ContainSearch() 函数实现包含关系查询的功能,输入一个面状的空间对象,返回该空间对象所包含的所有空间对象

• CrossSearch() 函数实现穿越查询的功能,输入任意一条弧线或折线,返回该弧线或折线所穿越的空间对象

• WithinSearch() 函数实现落入查询的功能,输入任意一个空间对象,如果该对象落入其它空间对象则返回所落入的空间对象,否则返回假值

• BufferSearch() 函数实现缓冲查询的功能,输入要缓冲查询的空间对象和缓冲条件,返回生成的缓冲区内所有空间对象。

4. 缓冲区分析组件 缓冲区分析的概念与缓冲区查询的概念完全不同,缓冲区查询是不破坏原有空间目标的关系,只

是检索得到该缓冲区范围内涉及到的空间目标。缓冲区分析则不同,它是对一组或一类地物按缓冲的距离条件,建立缓冲区多边形图,然后将这一个图层与需要进行缓冲区分析的图层进行叠置分析,得到所需要的结果。缓冲区分析包括点缓冲区、线缓冲区和面缓冲区。所以该组件需要向外界提供一个接口:IBufferAnalyzer(缓冲分析接口)调用 Point-Buffer() (点缓冲)、Polyline-Buffer() (线缓冲)、Polygon-Buffer() (面缓冲)这三种缓冲区分析的函数。以下是各个方法的描述:

• Point-Buffer() 函数实现点的缓冲区分析功能,输入一个点和缓冲的条件(即缓冲半径),返回缓冲区多边形同需要进行缓冲分析的图层的叠置分析的结果。

• Polyline-Buffer() 函数实现线的缓冲区分析功能,输入一条折线和缓冲半径,返回缓冲区多边形同需要进行缓冲分析的图层的叠置分析的结果。

• Polygon-Buffer() 函数实现多边形的缓冲分析功能,输入一个多边形和缓冲半径,返回缓冲区多边形同需要进行缓冲分析的图的叠置分析的结果。

5. 多边形分析组件 应该称为叠置分析组件,因为实际上多边形分析组件主要是考虑多边形同点、线和面的叠加关系。叠置分析是 GIS 中的一项非常重要的空间分析功能。它主要包括点同多边形的叠置分析(即多边形对点的包含关系)、线同多边形的叠置分析以及多边形的叠加分析。其中多边形的叠置分析最为复杂,它包括了多边形的交、并、差和异或等关系。所以该组件需要一个接口:IPolygonOperation(多边形分析接口),以及 Point-In-Polygon() (点同多边形叠置)、Polyline-In-Polygon() (线同多边形叠置)、Polygon-UNION() (多边形的并)、Polygon-DIFF() (多边形的差)、Polygon-INT() (多边形的交)、Polygon-XOR() (多边形异或)等叠置分析的方法。以下是各个方法的描述:

• Point-In-Polygon() 函数实现点对多边形的叠置分析功能,输入一个点和一个多边形,如果该点在多边形内返回真,否则返回假

• Polyline-In-Polygon() 函数实现线对多边形的叠置分析功能,输入一个多边形和一条折线,如果该折线完全落入多边形内返回真,否则返回假

• Polygon-UNION() 函数实现对两个多边形的并的功能,输入两个多边形,如果存在并集,返回合并后的多边形,否则返回假

• Polygon-DIFF() 函数实现对两个多边形的求差的功能,输入两个多边形,返回第一个多边形减去第二个多边形后得到的新的多边形

• Polygon-INT() 函数实现对两个多边形求交的功能,输入两个多边形,如果存在交集,则返回交集得到的多边形,否则返回假

• Polygon-XOR() 函数实现对两个多边形求异或的功能,输入两个多边形,返回两个多边形经过异或运算得到的多边形。

6. 网络分析组件 网络分析是运筹学模型中的一个基本模型,它的根本目的是研究、筹划一项网络工程如何安排,并使其运行效果最好,如一定资源的最佳分配,从一地到另一地的运输费用最低等。这类问题在社会经济活动中不胜枚举,因此在地理信息系统中此类问题的研究具有重要意义。根据目前实际的需要我们暂时主要实现以下功能:网络最短路径、累计属性值、邮路分配、空间邻域搜索、最近邻域搜索、地址匹配以及动态分割。所以需要有一个接口:INetAnalyer(网络分析接口)以及 Shortcut() (最短路径)、AccumulativeTotal() (累计

属性值)、PostDispatch() (邮路分配)、NeighborSearch() (邻域搜索)、NearestAdjacentSearch() (最近邻域搜索)、AddressMatching() (地址匹配) 和 DynamicDivision() (动态分割) 等多个方法实现该组件。以下是各个方法的描述:

• Shortcut() 函数实现最短路径的功能, 输入两个网络节点, 返回它们之间的最短路径

• AccumulativeTotal() 函数实现对属性值累计的功能, 输入某个属性, 返回该属性值的累积总和

• PostDispatch() 函数实现对邮路分配的功能, 输入邮件到达的网络节点, 返回邮路分配的最优路径

• NeighborSearch() 函数实现邻域搜索的功能, 输入某个空间对象, 返回它的邻域对象

• NearestAdjacentSearch() 函数实现最近邻域搜索功能, 输入某个空间对象, 返回离它最近的邻域

• AddressMatching() 函数实现地址匹配功能, 输入某个空间对象, 返回该对象所代表的现实对象的地址

• DynamicDivision() 函数实现动态分割的功能, 输入需要分割的图层和分割条件, 返回分割后得到的各个部分。

7. 地图分析组件 地图分析并不是 GIS 中所特有的功能, 在其它的一些图形软件中也有类似的功能, 它主要包括图层叠加、加减地图、形态分析、重新编码与分类、平均格网值、最大最小格网值、逻辑组合、乘除地图、窗口内平均值、窗口内最大最小值、窗口内总值以及窗口内最频繁格网值等功能。它需要一个接口 IMapAnalyzer() (地图分析组件接口) 和 LayerAdd() (图层叠加)、AddDiffMap() (加减地图)、ShapeAnalyzer() (形态分析)、Recoding() (重新编码与分配)、AverageMeshValue() (平均格网值)、MaxMinMesh() (最大最小格网值)、LogicCorporation() (逻辑组合)、MultipleDivideMap() (乘除地图)、Window Average Value() (窗口内平均值)、Window MaxMinValue() (窗口内最大最小值)、WindowTotalValue() (窗口内总值)、MostFrequentMesh() (窗口内最频繁格网值) 等方法来实现该组件。以下是各个方法的描述:

• LayerAdd() 函数实现图层叠加的功能, 输入图层和要叠加的对象, 返回叠加了该对象后的图层

• AddDiffMap() 函数实现图层的加减功能, 输入两个图层和操作符, 根据操作符返回两个图层相加或相减后得到的图层

• ShapeAnalyzer() 函数实现形态分析的功能, 输入要分析的图层, 返回分析后的形态报告

• Recoding() 函数实现重新编码与分配的功能, 输入要重新编码和分配的图层, 返回重新编码和分配后的图层

• AverageMeshValue() 函数实现平均格网值的计算功能, 输入一个图层, 返回平均格网值计算的结果

• MaxMinMesh() 函数实现最大最小格网值的计算功能, 输入一个图层, 返回最大和最小格网值

• LogicCorporation() 函数实现逻辑组合功能, 输入要组合的对象和组合条件, 返回按条件组合后的对象

• MultipleDivideMap() 函数实现地图的乘除功能, 输入一个图层和操作符, 以及一个给定的值, 返回根据操作符进行相乘或相除后得到的放大或缩小的图层

• WindowAverageValue() 函数实现窗口内平均值的计算, 输入一个窗口和要计算的属性, 返回窗口内该属性的平均值

• WindowMaxMinValue() 函数实现窗口内最大最小值的计算, 输入一个窗口和某个属性, 返回窗口内该属性的最大和

最小值

• WindowTotalValue() 函数实现窗口内总值的计算, 输入一个窗口和某个属性, 返回窗口内该属性的总值

• MostFrequentMesh() 函数实现窗口内最频繁格网值的计算, 输入一个窗口, 返回该窗口内最频繁的格网的值。

8. 地形分析组件 地形分析是一个复杂的组件, 它包括了坡度计算、任意点内插高程、点通视分析、线或面通视分析、生成等高线、生成断面以及计算最优路径等的分析。地形分析被广泛地应用在地理信息系统中的洪水险情预报分析、土地利用现状的分析等重要的应用中。它需要定义一个 ITerrainAnalyzer (地形分析) 接口以及 Gradient() (坡度分析)、InsertAltitude() (内插高程)、Point-Intervisibility() (点通视分析)、Polyline-Intervisibility() (线通视分析)、Area-Intervisibility() (面通视分析)、Contour() (生成等高线)、Geosutures() (生成断裂线)、OptimalRoute() (最优路径)、FillUp() (充填方分析) 等方法来实现各种功能。以下是各个方法的描述:

• Gradient() 函数实现坡度的计算, 返回一个坡的坡度

• InsertAltitude() 函数实现内插高程的功能, 输入要内插的值, 返回内插高程后得到的图形

• Point-Intervisibility() 函数实现点通视的分析, 输入一个被测试的点和测试源点, 如果两点之间存在通视 (即可见, 没有障碍物) 则返回真, 否则返回假

• Polyline-Intervisibility() 函数实现线通视的分析, 输入一条被测试的折线和测试源点, 如果整条折线同源点都通视则返回真, 否则返回假

• Area-Intervisibility() 函数实现面通视的分析, 输入一个被测试的面和一个测试源点, 如果整个面同测试源点都通视则返回真, 否则返回假

• Contour() 函数实现等高线的计算, 输入一个图层, 返回计算后得到的等高线

• Geosutures() 函数实现断裂线的计算功能, 根据输入图层, 返回一条根据用户定义所计算得到的断裂线

• OptimalRoute() 函数实现最优路径的分析功能, 输入判断的条件和两点坐标, 返回两点间最优路径

• FillUp() 函数实现充填方的分析功能, 输入要计算的部分的地面多边形和高度曲面的函数, 返回充填方的值 (即体积)。

结束语 组件式软件技术已经成为当今软件技术的潮流之一, 为了适应这种技术潮流, “GIS 软件象其他软件一样, 已经或正在发生着革命性的变化, 即由过去厂家提供了全部系统或者具有二次开发功能的软件, 过渡到提供组件由用户自己再开发的方向上来”。我国 GIS 发展了近二十年, 但目前仍未形成有规模的产业。组件式 GIS 开发平台的出现是推动我国 GIS 软件产业和应用事业的一个重要机遇。而本文提出的空间分析组件正是在组件技术的基础上开发的 GIS 中空间分析这一重要功能的软件。希望它的开发能够为其它 GIS 组件的开发提供技术上的参考和借鉴, 促进我国 GIS 组件化的发展。

参考文献

- 1 陈述彭, 等. 地理信息系统导论. 科学出版社, 2000
- 2 郭伦, 等. 地理信息系统——原理、方法和应用. 科学出版社, 2001
- 3 龚健雅. 地理信息系统基础. 科学出版社, 2001
- 4 钟耳顺, 宋关福, 王尔琪, 吴秋华. 地理信息系统论坛, 2000
- 5 Dale Rogerson. COM 技术内幕. 清华大学出版社, 1997