

地理信息系统概论

1.1 地理信息系统

1.1.1 信息和地理信息

1. 信息

信息(information)是用数字、文字、符号、语言等介质来表示事件、事物、现象等的内容、数量或特征,信息向人们(或系统)提供关于现实世界新的事实知识,作为生产、管理、经营、分析和决策的依据。

信息具有客观性、适用性、可传输性和共享性等特征。客观性是指信息都与客观事实相关,这是信息正确性和精确度的保证;适用性是指从大量数据中收集、组织和管理有用的信息,这是由建立信息系统的目的性所决定的;可传输性是指信息可以在系统内或用户之间以一定形式或格式进行传送和交换,随着网络技术的发展,信息的可传输性日益重要;共享性是信息可传输性带来的结果,也就是信息可为多个用户共享。

信息来自数据,数据是未加工的原始资料。数字、文字、符号、图形和影像都是数据,数据是客观对象的表示。信息则是数据内涵的意义,是数据的内容和解释。例如,从测量数据中可以抽取出目标和物体的形状、大小和位置等信息,从遥感卫星图像数据中可以抽取出各种图形和专题信息,从实地调查数据中则可抽取出各专题的属性信息。

2. 地理信息

地理信息(geographic information)是指与所研究对象的空间地理分布有关的信息,它表示地表物体和环境固有的数据、质量、分布特征,是具有联系和规律的数字、文字、图形、图像等的总称。

地理信息属于空间信息,与一般信息的区别在于它具有区域性、多维性和动态性。区域性是指地理信息的定位特征,且这种定位特征是通过公共的地理基础来体现的。例如,用经

纬网坐标或公里网坐标来识别空间位置,并指定特定的区域。多维性是指在一个坐标位置上具有多个专题和属性信息。例如,在一个地面点上可取得高程、污染、交通等多种信息。动态性是指地理信息的动态变化特征,即时序特性。这一特性使地理信息常以时间尺度划分成不同时间段的信息,这就要求及时采集和更新地理信息,并根据多时相数据和信息来寻找时间分布规律,进而对未来作出预测和预报。

客观世界是一个庞大的信息源,随着现代科学技术的发展,特别是借助近代数学、空间科学和计算机科学,人们已能够迅速地采集到地理空间的几何信息、物理信息和人文信息,并适时适地地识别、转换、存储、传输、显示并应用这些信息,进一步为人类服务。

1.1.2 信息系统和地理信息系统

1. 信息系统

信息系统(information system)是具有采集、处理、管理和分析数据能力的系统,它能为单一的或有组织的决策过程提供各种有用信息,其功能框图见图 1.1。信息系统的四大功能为数据采集、管理、分析和表达。更简单地说,信息系统是基于数据库的问答系统(见图 1.2)。

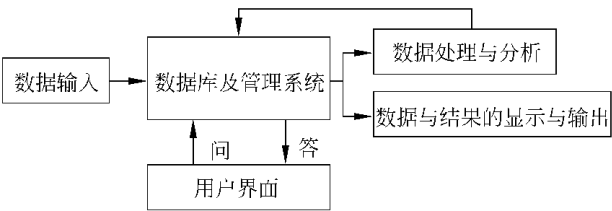


图 1.1 信息系统功能框图

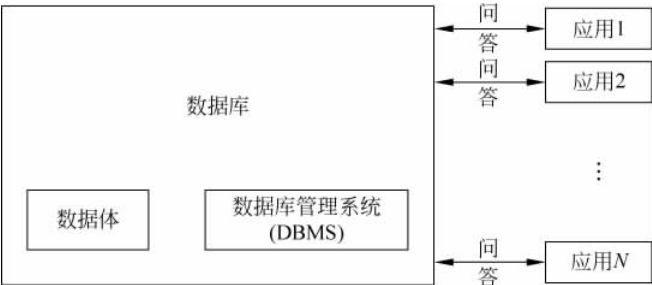


图 1.2 作为问答系统的信息系统

从计算机科学角度看,信息系统是由计算机硬件、软件、数据和用户四大要素组成的问答系统,智能化的系统还包括知识(见图 1.3)。硬件包括各类计算机处理器及其终端设备;软件是支持数据与信息的采集、存储、加工、再现和回答用户问题的计算机程序系统;数据则包括定量数据和定性数据;用户是信息系统所服务的对象,是信息系统的主人。

用户包括一般用户和从事系统建立、维护、管理和更新的高级用户。

由于计算机技术的飞速发展以及计算机应用的普及,不同问题领域的各种信息系统相

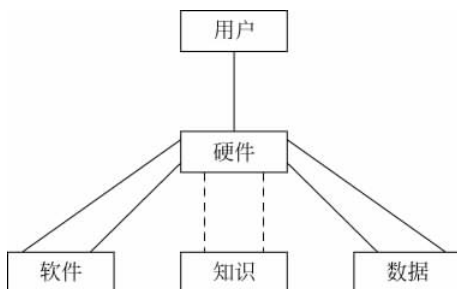


图 1.3 计算机科学意义上的信息系统

继出现,且种类繁多,但从系统结构及处理方法看,主要分为以下几种。

(1) 管理信息系统(management information system,MIS)。它是一种基于数据库的问答系统,它往往停留在数据级上支持管理者,如人事管理信息系统、财务管理信息系统、产品销售信息系统等。

(2) 决策支持系统(decision support system,DSS)。它是在 MIS 基础上发展起来的一种信息系统,它不仅为管理者提供数据支持,还提供方法和模型的可能支持,并对问题进行仿真和模拟,从而辅助决策者进行决策。

(3) 智能决策支持系统(intelligent decision support system,IDSS)。它是在决策支持系统中进一步引入人工智能(artificial intelligence,AI)技术,如专家系统(expert system,ES)来解决非结构化问题,提高系统决策的自动化程度。

(4) 空间信息系统(spatial information system,SIS)。它是对空间数据进行采集、处理、管理和分析的信息系统。由于空间数据的特殊性,空间信息系统的组织结构及处理方法有别于一般信息系统。

2. 地理信息系统

地理信息系统(GIS)是一种特定而又十分重要的空间信息系统,它是用于采集、存储、管理、分析和描述整个或部分地球表面(包括大气层在内)与空间和地理分布有关的数据的空间信息系统。由于地球是人们赖以生存的基础,所以 GIS 是与人类的生存、发展和进步密切相关的一门信息科学与技术,受到人们越来越广泛的重视。

随着地理信息系统的广泛应用,其在环境科学领域也相继产生了多种系统,如自然资源管理信息系统(natural resources management information system)、资源与环境信息系统(resources and environment information system)、土地资源信息系统(land resources information system)、空间数据处理系统(spatial data processing system)、空间信息系统(spatial information system)。虽然这些系统的研究对象不同,但研究方法却是基本相似的。

地理信息系统是一门多技术交叉的空间信息科学,它依赖于地理学、测绘学、统计学等基础性学科,又与计算机硬件与软件技术、航天技术、遥感技术和人工智能与专家系统技术的进步与成就息息相关,如图 1.4 所示。此外,地理信息系统还是一门以应用为目的的信息产业,它的应用可深入到各行各业。

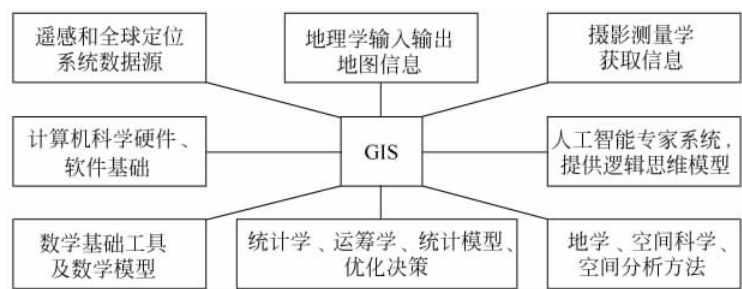


图 1.4 地理信息系统与相关学科

1.1.3 地理信息系统的类型

地理信息系统应用面广,技术潜力大,且发展极为迅速,因此很难用一种固定方法进行分类。但是,通常可从下面几个角度来分类。

1. 以研究对象性质和内容分类

(1) 综合性地理信息系统。按国家统一标准,存储管理全国范围内的各种自然和社会经济数据的地理信息系统,或对全球气候、人口、资源进行存储管理的全球地理信息系统,如加拿大国家地理信息系统、中国自然环境综合信息系统等。

(2) 专题性地理信息系统。指以某一专业、任务或现象为目标建立的地理信息系统,这种系统中数据项的内容及操作功能的设计都是为某一特定专业任务服务的,如小流域综合治理地理信息系统、森林资源管理信息系统等。

2. 以研究对象分布范围分类

(1) 全球性地理信息系统。这种系统研究区域范围往往涉及全球范围,如全球人口资源地理信息系统。

(2) 区域性地理信息系统。指以某个区域(如行政区)为对象进行研究管理、规划的信息系统,如美国明尼苏达州土地管理信息系统、我国黄土高原地理信息系统等。

3. 以地理信息系统应用功能分类

(1) 工具型地理信息系统。地理信息系统是一个复杂庞大的空间管理信息系统。用地理信息系统技术实际问题时有大量软件开发任务,如果用户重复开发,会对人力、财力等造成巨大的浪费。工具型地理信息系统为地理信息系统的使用者提供一种技术支持,使用户能借助地理信息系统工具中的功能直接完成应用任务,或者利用工具型地理信息系统加上专题模型完成应用任务。目前,在国内外已有很多商业化的工具型地理信息系统,如美国 ESRI 公司的 Arc GIS、美国 MapInfo 公司的 MapInfo、中国武汉中地公司的 MapGIS、中国北京超图公司的 SuperMap、中国武汉吉奥公司的 GeoStar 等。此外,近些年来,很多基于开源(open source)软件的工具型地理信息系统也得到了迅速发展并日益成熟,如 PostGIS、Grass、MySQL、OpenLayers、Mapnik、GDAL 等。目前,这些开源的软件产品正受到越来越

多的 GIS 工作者的重视与青睐,其应用也变得越来越广泛。

(2) 应用型地理信息系统。应用型地理信息系统的开发分两类:一类是借助工具型地理信息系统开发的;另一类是为某专业部门应用自行开发的,这种系统针对性明确、专业性强、系统开销小,适于在本专业中推广使用。

4. 以地理信息系统数据结构类型分类

地理信息系统的数据结构直接影响它的输入、存储、管理及输出方法和手段。

(1) 矢量数据结构地理信息系统。它是指以 x 、 y 坐标对来表示空间数据的点、线和面等图形的地理信息系统。

(2) 栅格数据结构地理信息系统。它是指以二维数组来表示空间各像元特征的地理信息系统。

(3) 混合数据结构地理信息系统。由于矢量数据结构地理信息系统和栅格数据结构地理信息系统的特点不同,适用范围不同,相互之间不能互相替代,因此出现了矢量数据结构和栅格数据结构并存的地理信息系统。

矢量数据和栅格数据结合通常采用矢量数据和栅格数据相互间转换来实现。这方面出现了很多算法,但这种转换不仅花费时间,而且经过转换后的原始信息会受到不同程度的损失。为此,研究者一直在探索和寻找一体化的数据结构,即混合数据结构。同时,在这方面也出现了很多自动的计算机算法,取得了一些成绩并日臻成熟起来。

1.2 地理信息系统的结构与功能

1.2.1 地理信息系统的组成

地理信息系统主要由 4 部分组成:计算机硬件系统,计算机软件系统,空间数据及系统使用、管理和维护人员(即用户)。地理信息系统的核心内容是计算机硬件和软件,空间数据反映了应用地理信息系统的信息内容,用户决定了系统的工作方式,如图 1.5 所示。

1. 计算机硬件系统

计算机硬件系统是计算机系统中实际物理设备的总称,主要包括计算机主机、输入设备、存储设备和输出设备。

2. 计算机软件系统

计算机软件系统是地理信息系统运行时所必需的各种程序,具体内容如下。

(1) 计算机系统软件。

(2) 地理信息系统软件及其支撑软件。包括地理信息系统工具或地理信息系统实用软件程序,以完成空间数据的输入、存储、转换、输出及用户接口功能等。

(3) 应用程序。这是根据专题分析模型编制的具有特定应用任务的程序,是地理信息系统功能的扩充和延伸。一个优秀的地理信息系统工具,对应用程序的开发应是透明的。

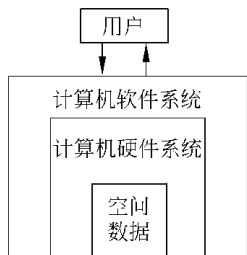


图 1.5 地理信息系统的构成

应用程序作用于专题数据上,构成专题地理信息系统的基本内容。

3. 空间数据

空间数据是地理信息系统的重要组成部分,是系统分析加工的对象,是地理信息系统表达现实世界的、经过抽象的实质性内容。它一般包括 3 个方面的内容,即空间位置坐标数据、地理实体之间的空间拓扑关系以及相应于空间位置的属性数据。通常,它们以一定的逻辑结构存放在空间数据库中,空间数据来源比较复杂,根据研究对象不同、范围不同、类型不同,可采用不同的空间数据结构和编码方法,其目的就是为了更好地管理和分析空间数据。

4. 系统使用、管理和维护人员

地理信息系统是一个复杂的系统,仅有计算机硬件、软件及数据还不能构成一个完整的系统,必须要有系统的使用管理人员,包括具有地理信息系统知识和专业知识的高级应用人才、具有计算机知识和专业知识的软件应用人才以及具有较强实际操作能力的软硬件维护人才。

1.2.2 地理信息系统的主要功能模块

地理信息系统软件一般由 5 部分组成,即由空间数据输入管理、空间数据库管理、应用模型、空间数据处理和分析及空间数据输出管理组成,它们之间的关系见图 1.6。

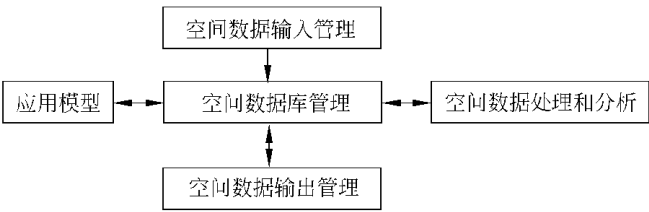


图 1.6 地理信息系统的主要功能模块

1. 空间数据输入管理

空间数据输入管理模块是相对独立的功能模块,它的目的是将地理信息系统中各种数据源输入,并转换成计算机所要求的数据格式进行存储。随着数据源种类的不同(如文本数据、数字数据和模拟数据等)、输入设备的不同及系统选用的数据结构及数据编码的不同,在数据输入部分配有不同的软件,以确保原始数据按要求存入空间数据库中。

通常,空间数据输入的同时,伴随着对输入数据的处理,以实现数据的校验和编辑。

2. 空间数据库管理

与一般数据库相比,地理信息系统数据库不仅要管理属性数据,还要管理大量图形数据,以描述空间位置分布及拓扑关系。另外,属性数据和图形数据之间具有不可分割的联系。

此外,地理信息系统中数据库的数据量大,涉及内容多,这些特点决定了它既要遵循常

用关系型数据库管理系统来管理数据,又要采用一些特殊的技术和方法,来解决通常数据库没法管理的空间数据问题。由于地理信息系统数据库具有明显的空间性,所以也称其为空间数据库,其组成如图 1.7 所示。

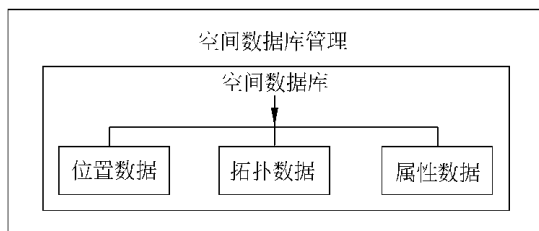


图 1.7 空间数据库管理

3. 应用模型

由于地理信息系统应用范围越来越广,常规系统提供的处理和分析功能很难满足所有用户的要求,因此一个优秀的地理信息系统应当为用户提供二次开发手段,以使用户开发新的空间分析模块,即开发各种应用模型,扩充地理信息系统功能。

4. 空间数据处理和分析

空间数据处理和分析模块通常为地理信息系统提供一些基本的和常用的处理和分析功能,其功能的强弱直接影响地理信息系统的应用范围,因此这部分是体现地理信息系统功能强弱的关键部分。

5. 空间数据输出管理

地理信息系统中输出数据种类很多,可能是输出地图、表格、文字、图像等;输出介质可以是纸、光盘、磁盘、显示终端等。随着输出数据类型的不同和输出介质的不同需配备不同的软件,最终向用户报告分析结果。

1.3 地理信息系统的特性

如上所述,计算机制图、计算机辅助设计、数据库管理系统、遥感图像处理技术奠定了地理信息系统的技术基础。地理信息系统是这些学科的综合,它与这些学科和系统之间既有联系又有区别,这里将它们逐一加以比较,以突出地理信息系统的特点。

1.3.1 GIS 与机助制图系统的区别与联系

机助制图是地理信息系统的主要技术基础,它涉及 GIS 中的空间数据采集、表示、处理、可视化甚至空间数据的管理。无论是在国外还是在国内,GIS 早期的技术都主要反映在机助制图方面。机助制图系统或者说数字地图系统涵盖了相当大的范围,从大比例尺的数字测图系统、电子平板,到小比例尺的地图编辑出版系统、专题图的桌面制图系统、电子地图

制作系统及地图数据库系统。它们的功能主要强调空间数据的处理、显示与表达,有些数字制图系统包含空间查询功能。

地理信息系统和数字地图制图系统的主要区别在于空间分析方面。一个功能完善的地理信息系统可以包含数字地图制图系统的所有功能,此外它还应具有丰富的空间分析功能。

尽管利用数字地图的集合可以建立一个数字地图库,并用数据库管理技术对其实现查询和检索功能,但它决不可能像地理信息系统那样提供出规划和决策方案。这是因为地理信息系统往往根据不同专业要求配有相应的分析模型,因此它有很强的处理分析能力。

实际上,数字地图及制图应该是地理信息系统的重要组成部分。首先,数字地图是地理信息系统重要的数据源,数字地图制图系统中存储和管理的信息往往是地理信息系统所需要的。其次,地理信息系统中,处理分析结果常以数字地图的形式来表现和输出。例如,对某区域进行土地利用规划后输出土地利用规划图,该输出功能包括数字地图的制图。

1.3.2 GIS 与数据库管理系统的区别与联系

目前,使用最为普遍的数据库管理系统是关系型数据库管理系统,如 Oracle、SyBase、SQL Sever、Infomix、FoxPro 等。它们不仅是一般事务管理系统,如银行系统、财务系统、商业管理系统、飞机订票系统等的基础软件,而且通常也是地理信息系统中属性数据管理的基础软件。目前,甚至有些 GIS 的图形数据也交给关系型数据库管理系统管理,而关系型数据库管理系统也在向空间数据管理方面扩展,如 Oracle、Infomix、Ingres 等都增加了管理空间数据的功能,用来对 GIS 中的图形数据和属性数据进行一体化的管理。

但是数据库管理系统和地理信息系统之间还存在着区别。地理信息系统除需要功能强大的空间数据的管理功能之外,还需要具有图形数据的采集、空间数据的可视化和空间分析等功能。所以,GIS 在硬件和软件方面均比一般事务数据库更加复杂,在功能上也比后者要多得多。例如,电话查号台可看作一个事务数据库系统,它只能回答用户所查询的电话号码,而一个用于通信的地理信息系统除了可查询电话号码外,还可提供所有电话用户的地理分布、电话空间分布密度、公共电话的位置与分布、新用户与最近电信局距离等信息。

1.3.3 GIS 与 CAD 的区别与联系

计算机辅助设计(CAD)是计算机技术应用于机械、建筑、工程和产品设计的系统,它主要用于范围广泛的各種产品和工程的图形的辅助设计,大至飞机小到微芯片等。

GIS 与 CAD 系统的共同特点是两者都有坐标参考系统,都能描述和处理图形数据及其空间关系,也都能处理非图形属性数据。它们的主要区别是:CAD 处理的多为规则几何图形及其组合,图形功能极强,属性功能相对较弱;而 GIS 处理的多为地理空间的自然目标和人工目标,图形关系复杂,需要有丰富的符号库和属性库。GIS 需要有较强的空间分析功能,图形与属性的相互操作十分频繁,且多具有专业化的特征。此外,CAD 一般仅在单幅图上操作,海量数据的图库管理的能力比 GIS 要弱。

但是由于 CAD 具有极强的图形处理能力,也可以设计丰富的符号和连接属性,许多用户都把它作为数字制图系统使用。有些地理信息系统将 CAD 作为数据采集的辅助工具。例如 AutoCAD 软件,与不同 GIS 软件系统之间有接口,可以把 AutoCAD 输入的图形数据

传送给地理信息系统。美国 ESRI 公司和 AutoDESK 公司合作推出的 ARC. CAD, 可以同 AUTOCAD 一起在微机上实现地理信息系统功能, 并同 GIS 软件 Arc/Info 有机地结合起来。

1.3.4 GIS 与遥感图像处理系统的区别与联系

遥感图像处理系统是专门用于对遥感图像数据进行分析处理的软件。它主要强调对遥感栅格数据的几何处理、灰度处理和专题信息提取。遥感数据是地理信息系统的重要信息源。遥感数据经过遥感图像处理系统处理之后, 或是进入 GIS 系统作为背景影像, 或是与经过分类的专题信息一道协同进行 GIS 与遥感图像的集成分析。

一般来说, 遥感图像处理系统还不便于用作地理信息系统。然而, 许多遥感图像处理系统的制图功能较强, 可以设计丰富的符号和注记, 并可进行图幅整饰, 生产精美的专题地图。有些基于栅格的 GIS 除了能进行遥感图像处理之外, 还具有空间叠置分析等 GIS 的分析功能。但是这种系统一般缺少实体的空间关系描述, 难以进行某一实体的属性查询和空间关系查询以及网络分析等功能。当前遥感图像处理系统和地理信息系统的发展趋势是两者的进一步集成, 甚至研究开发出在同一用户界面内进行图像和图形处理以及实现矢量、栅格影像和 DEM 数据的整体结合的存储方式。

1.3.5 地理信息系统和事务处理系统的区别与联系

地理信息系统和一般事务处理系统及信息系统的主要区别在于地理信息系统处理的数据是空间数据, 它不仅管理反映空间属性的一般的数字、文字数据, 还要管理反映地理分布特征及其之间拓扑关系的空间位置数据, 而且还要把两者有机结合起来进行协调管理和分析, 而事务处理系统相对简单得多。此外, 地理信息系统对计算机硬件和软件资源的要求比一般事务处理系统高。例如, 地理信息系统必须具有处理空间数据的输入、输出装置, 如数字化仪、扫描仪、绘图仪等。再有, 由于地理信息系统处理数据量大、运算复杂, 对计算机存储量、运算速度等的要求相应也高。

1.4 地理信息系统的发展与前景

地理信息系统是一类获取、处理、分析、表示并在不同系统、不同地点和不同用户之间传输空间数据的计算机应用系统。由于地球上 80% 以上的信息与空间位置有关, 它应该成为在操作系统、数据库管理系统之上的主要应用集成平台, 占据软件产品的主流地位。但是, 目前 GIS 软件在技术上并不能适应飞速发展的应用要求, 尤其不能适应“数字地球”“数字城市”和“数字区域”建设中数据集成和功能集成的需要。认真研究现有 GIS 理论与软件实现技术的不足, 寻求理论、方法、技术和工具 4 个层次的突破和创新, 已经成为研究的热点。

1.4.1 地理信息系统软件发展的历史回顾

自从 1960 年加拿大测量学家 Roger F. Tomlinson 提出“要把地图变成数字形式的地图, 便于计算机处理与分析”的观点以来, 它就一直是研究与发展 GIS 软件的指导思想。纵

观 GIS 50 多年的发展历史, GIS 软件技术及其应用取得了巨大的发展,但也存在着严重的不足。从技术层面看,其发展大致可以分成 4 个阶段。

第一阶段是从 20 世纪 60 年代到 80 年代中后期,是 GIS 软件从无到有、从原型到产品的阶段。由于各种条件,包括自身理论和实现技术的不成熟和 IT 技术的限制,这一阶段的 GIS 软件存在许多不足。第一代 GIS 的基本技术特点如下。

(1) 以图层作为处理的基础。由于当时 GIS 系统中空间数据的主要来源是纸质地图的数字化, GIS 的数据模型中图层处于中心地位。定义某一区域轮廓组成图层,同一图层的空间数据存放在一个文件之中。空间定位与量算以图层范围内的平面笛卡儿坐标系为基础,操作局限于当前图层内。利用计算机技术可以实现同一区域内各种专题数据的叠置、影响区域分析(缓冲)和线状实体的路径分析。但是,各类查询与计算只能在同一图层中进行。

(2) 以系统为中心。当时的 GIS 系统中的空间数据各自有自己的数据格式,自成系统,不同的 GIS 系统基本上没有联系。与其他的软件工具,如 MRPII 软件、CAD 软件等基本上不存在集成关系。

(3) 单机、单用户。由于 IT 技术的限制,当时的 GIS 系统只能如此。

(4) 全封闭结构,支持二次开发的能力非常弱。当时的 GIS 只提供功能极其有限的查询和计算的自定义语言,与数据库、通用的编程语言没有建立联系,用户只能按照已经开发好的应用系统功能接口操作,或者联机地、一步一步地完成自己预定的计算任务,无法连贯地、批量地实现复杂的自定义应用功能。

(5) 在主要实现技术上,以文件系统来管理空间数据与属性数据。GIS 的数据分为空间数据和属性数据两类,空间数据描述空间实体的地理位置及其形状,属性数据则描述相应空间实体有关的应用信息。由于当时数据库管理系统只能管理结构化数据,对空间数据这样的非结构化数据无法进行定义、管理与操纵, GIS 软件只能在文件系统中自行定义空间数据结构及其操纵工具。由于最初关系型 DBMS 不够成熟与普及,对属性数据这样的结构化数据也放在文件系统中进行管理。

(6) 应用领域基本上集中在资源与环境领域的管理类应用。

第二阶段从 20 世纪 80 年代末到 90 年代中后期,是 GIS 软件逐渐成熟并广泛应用的阶段。在这一阶段, GIS 软件作为一种软件工具,理论与技术已经基本成熟。由于其具备空间数据操纵能力,在应用中受到青睐,应用领域迅速扩展。巨大的应用前景对 GIS 软件提出了各种各样的要求, GIS 软件实现技术也得到迅速发展,但是 GIS 的基本技术体系仍然没有发生根本的变化。第二代 GIS 的基本技术特点如下。

(1) 以图层作为处理基础。由于空间数据模型没有根本的变化,以图层为处理基础的模式依然没有变化。对属性数据的查询可以跨图层进行,但对空间数据的操作仍然限制在同一图层之内。在 GIS 应用系统事先定义的应用功能以外,联机的步进操作方式没有变化。

(2) 引入网络技术,实现了多机、多用户模式。由于这一阶段网络技术已经成熟,应用范围迅速扩大, GIS 软件也转向多用户和客户机/服务器(client/server)结构。但是,由于数据模型没有根本变化,客户机与服务器的关系基本上属于空间数据文件下载和回送的关系,基本的空间数据处理功能在客户端实现,因而是一种典型的“胖客户”类型。服务器只作文件服务器用,以 NFS 技术为基础,是一种典型的 C/S 结构。