

# 从数字地图到空间信息网格 ——空间信息多级网格理论思考

李德仁 朱欣焰 龚健雅

测绘遥感信息工程国家重点实验室 湖北省武汉市珞瑜路 129 号 430079

摘要：本文从思考地理空间数据在计算机网络系统中的表示方法出发，分析了网络计算机环境对 GIS 提出的挑战。在较详细回顾和总结了近二十年来在地理空间信息技术方面已取得的主要成果和存在问题的基础上，作者提出了一种既能适合网格计算环境又充分考虑到地球空间的自然特征和社会属性的差异性及其经济发展不平衡的特点的空间信息表示新方法——空间信息多级网格。

文中进而介绍和阐述了空间信息多级网格的体系结构、信息表示与数据存取、与传统各种比例 GIS 空间数据库之间的数据转换问题。文章对这些方面应深入开展研究的关键技术进行了阐明，并对空间信息多级网格技术在国家和省市宏观决策中的应用可能性进行了讨论。

关键词：网格计算 空间信息多级网格 多比例尺空间数据库 地理信息系统

## 1. 问题的提出

### 1.1. 数字地图是地理空间数据在计算机中表示的唯一方法吗？

自从发明了纸张之后，地理（地球）空间信息是以地图形式（地形图、专题图、透视/晕渲图和断面图等）在纸介质上表示的，已有几千年的历史。所以当电子计算机问世后，要把空间数据放到电子计算机去的时候，人们就自然地想到了“地图数字化”的道路。于是出现了用离散而且有拓扑关系的点串来描述点线面体各种空间要素。但是我们仔细地想一想，数字地图难道就是空间数据在计算机中表示的唯一方法吗？

不妨举几个例子来分析。当一个人拥有土地、房产时，他首先关心的是土地的面积和房屋的面积。人们可以用很简单的方法直接而精确地量出这些面积。但要从数字地图和计算机空间数据库中得到这些面积值，必须由描绘土地和房屋边界地点串坐标用计算方法求得。而且坐标系选择不同，这些点串就有不同的坐标。其实，面积并不会因坐标系的不同而发生变化。显然，用离散点串的坐标系描述地理空间要素的相对量并不是最好的方法。

另一个例子如人口普查。国家是想知道每个空间关系中人口的数量、年龄结构、文化结构、职业结构等信息。这要用地图来表示就得从 GIS 中去搜索每个空间关系中的住房数，从每个住房的属性库中查找上述信息并进行累加和分类。这种操作涉及到的计算对统计一个国家、一个洲和全球也是十分庞大的。如果当初人口调查时就直接按一个规定的网格单元来存贮这些数据，使用起来不是更方便吗？

所以不少的知名 GIS 学家都把地理空间数据在计算机中的表示方法列为 GIS 理论研究命题之一 [1][2]。

### 1.2. 网格计算环境对 GIS 的挑战

网格(grid)是近年来逐渐兴起的一个研究领域。当前的 Internet 技术实现了计算机硬件的连通,Web 技术实现了网页的连通,而网格技术是要把整个因特网上的各种资源整合成一台巨大的计算机,从而实现资源共享与协同工作[3]。用户分享网上资源,感觉如同个人使用一台超级计算机一样。网格技术将各种信息资源(内容)连接起来,比现有网络更有效的利用信息资源。

对网格的研究工作分为三个层次:计算网格、信息网格和知识网格[4]。计算网格是网格的系统层,它为应用层(信息网格、知识网格等)提供系统基础设施。信息网格研制一体化的智能信息处理平台,消除信息孤岛,使得用户能方便地发布、处理和获取信息。知识网格研制一体化的智能知识处理平台,消除知识孤岛,使得用户能方便地发布、处理和获取知识。

对比这样的网格计算环境,不难发现,现有的空间信息技术确实难以适应了,网格计算对 GIS 的挑战是显而易见的。

传统的空间信息系统没能很好地解决空间信息的共享和利用问题。

国家在过去几十年已经花费了大量的财力收集各种地下、地面、大气的各类资源数据，采用的手段包括地下勘探、地面人工测绘和监测、航空遥感、卫星遥感等。但目前由于零碎地分散在不同的地方，共享困难，利用率低。虽然地理信息系统在过去的十年里得到了很大的发展，在很多部门得到了广泛的应用，但总的说来，信息的共享程度不高，各个部门建立的与地理信息系统不能相互利用，形成信息“孤岛”。造成这种状况的原因包括政策方面和技术方面。在政策方面，国家还没有出台适宜的有关政策、法律和规章来有效地协调和指导整个国内空间信息发展，为全社会所充分使用，影响了空间信息共享的进程。在技术方面，还没有建立完善的空间信息共享标准体系，包括空间数据标准和互操作标准。

与一般的信息不一样，空间信息具有关系复杂、非结构化、数据量大、多比例尺、随时间变化等特点。在处理空间信息的拓扑关系、存储非结构化数据方面，传统的地理信息系统都已经比较成熟。对于海量数据处理，目前的 GIS 还有许多问题需要解决。然而，在实际应用中，由于不同行业、不同部门有不同内容的专业信息，空间数据的生产、维护都分散在不同的单位进行，很难进行全部统一集中。因此，比较可行的办法是各个行业、部门相对集中的分而治之，通过合理的技术途径，为其它部门提供服务。这正是网格技术要解决的问题。

空间数据种类繁多、数据格式各异、存在多种比例尺、多种空间参考系和多种投影类型。这给需要使用空间数据的用户带来了很大的困难，不利于空间信息的共享。然而，不同的应用需要不同比例尺的空间信息的支持，对空间参考系和投影类型也有相应的要求。目前的 GIS 数据大都来自地图数字化，而不是直接的测量数据。地图以某种统一空间基准测量得到的结果，经投影变换后地物在平面上表示，很难适应基准变化的要求。同时，由于测量技术、方法和设备仪器的限制，原来数据结果可能在不远的未来不能满足实际需要，现在的空间数据可能在未来的某个时候必须重建，而这种情况随着技术的进步依然会存在。现有的空间数据组织和管理技术没有很好的解决这些问题。因此很难适应未来空间信息共享和利用的要求。

要解决空间信息的共享问题，必须从根本上研究适合于空间信息共享的基础理论和方法，从空间信息的表示、数据的组织和管理、共享和服务模式上提出有别于目前空间信息系统的一套体系框架。网格技术与空间信息技术结合，有望解决共享和服务模式方面的问题，而适应空间信息网格的数据表示、组织方法则是需要研究的重点内容之一。

以上的种种问题，向我们提出的挑战是必须不失时机地开展空间信息网格 ( Spatial Information Grids ) 的研究，努力探求空间信息技术与网格技术的有机结合方法，更好地解决空间地理信息资源共享、分析应用与服务的问题。国家“十五”八六三计划，将它列为研究主题是完全适时而且十分必要的。

## 2. 地理空间信息技术目前已取得的成果与存在问题

为了深化对空间信息网格的认识和开展相应的研究，有必要对 GIS 问世的三十年来空间信息技术已取得的成果与存在的问题作扼要的回顾。

### 2.1. 地理信息系统

地理信息系统在过去的十年里得到了极大的发展。在 GIS 的应用体系结构发展上表现为：以主机 ( mainframe computer ) 为中心的 GIS 系统技术应用体系结构；基于桌面 PC 机 ( 微机 ) 为核心的 GIS 系统技术应用体系结构和当前刚刚兴起的支持 Internet/Intranet 网络技术标准的分布式 GIS 以及 Web 和 MobileGIS 系统技术应用体系结构。在数据管理上表现为[5]：文件系统方式；几何数据以文件管理方式，属性数据以关系数据库管理方式；全关系数据库方式，对象关系数据库方式。在开发方式上表现为：从专用环境下的专用开发包，发展为开放环境下的通用组件开发。在数据管理能力上，从小数据量管理，向海量数据管理发展。在应用领域上，从专业应用领域向公众信息发布、决策支持方面发展。

虽然地理信息系统得到了广泛的应用，但总的说来，空间信息的利用并不高，各个部门建立的与地理信息系统难以相互利用，形成信息“孤岛”。传统的空间信息系统没能很好解决空间信息的共享和利用问题。

究其原因，主要在于 GIS 是面向空间数据，特别是地图数据的，而不是直接面向空间信息的。空间信息必须用某种算法由空间数据导出来的。所以对用户特别对大型用户的大任务，它是不友好的，不方便的，有时也是十分费时的。

## 2.2. 多比例尺空间数据与信息的表示

地图是空间数据的一个重要数据源，GIS 最初也是由计算机辅助地图制图系统发展而来，而比例尺则是地图的一个基本而重要的概念。因此 GIS 也就不可避免地要表示各种不同比例尺的空间数据。不同比例尺的地图，表示了地面上地物目标的不同细节程度，同一地物目标在不同比例尺的地图上的表现形式也可能不同，如在大比例尺图上用面来表示的目标，在小比例尺上可能表现为点目标。然而，传统的模拟地图数字化变成了数字化产品后，在数字化世界里是否还需要比例这一概念成为一个有争议的问题。实际上，目前的数字化地图还不可能离开比例尺的概念，一方面，客观应用离不开不同比例尺的地图；另一方面，即使制图综合问题得到完全解决后，也不可能完全从最大比例尺的地图自动综合生成其他各种小比例尺的地图。也就是说，地图综合只能在一定比例尺范围内进行。因此，多比例尺空间数据的表示与处理是 GIS 的一个固有问题。

多比例尺的概念在影像和 DEM 数据中表现为不同的分辨率。用金字塔结构来组织不同的分辨率的影像和 DEM 数据已基本形成共识，这种表示对于数据的表现有较高的效率，可以实现不同分辨率数据之间的无缝浏览，在与矢量数据的一体化表现方面也具有优势。从某种程度上看，多比例尺的影像或 DEM 数据实现了某种程度上的互动。但影像数据和矢量数据之间，目前还不能自动从影像数据中提取矢量数据，还不具备互动能力。

对于矢量数据，多比例尺的数据组织与表示方法主要有三种[6]。

第一种是建立同一数据库的多个比例尺版本，如 1:5 万、1:25 万、1:50 万、1:100 万等各自建立一套数据库版本。这一方法的主要优点是简单，通过程序的方法可以实现多比例尺的之间跨越显示问题，可以实现不同比例尺的矢量数据之间的显示互动，但问题很多，如投入费用大，生产周期长，各数据库之间同一地理目标之间缺乏联系，导致数据更新和维护困难，自动更新几乎不可能实现，这种更新互动在目前的 GIS 中无能为力。

第二种方法主要是针对第一种方法的缺点而展开的研究。开发更好的数据模型来支持空间数据的多级表达，包括运用面向对象方法，语义数据模型等技术。如文献[7][8]提出的 SDS (Simplicial data structure) 数据结构，采用建立所有地物坐标点的详细三角形来表达地面，并允许包括移位和融合等操作的综合运算。这种方法着眼于用同一数据库表达同一地物的多个层次的比例尺版本，并建立内在的联系，试图解决不同比例尺数据版本之间的更新和维护互动问题。这种数据结构理论上可以分成任意层次，但实际上仍然是存在于有限的离散比例尺层次上。

第三种方法试图通过自动地图综合的方法来自动生成同一数据库的多个版本，虽然已经研究了很长时期，但仍未解决。地图综合的研究主要包括算法理论研究、模型和概念框架的设计、专家系统的应用以及制图特征的建模等几个方向[9]。迄今为止，大多数综合研究工作主要着重于与比例尺变化相关的“制图”综合的研究，而缺乏对“模型”综合的研究。对专家系统应用于自动综合研究也偏重于对知识的“获取途径”和“获取方法”的研究，而没有或很少考虑到从多比例尺或综合的角度对蕴含更多知识的空间数据库本身的组织结构的研究和对数据库建立过程的研究，很少研究考虑如何使空间数据库本身更好地支持地图综合或多比例尺表达与处理[10]。同时，对综合算法的研究虽然经历时间最长，研究成果最多，但一直难以很好地解决多比例尺表达的问题，对算法的研究仍然需要更多的努力。

如果从地图的束缚下解放出来，空间信息的表达、统计、分析是不一定需要，有时也不应当进行综合取舍的。例如统计全国森林面积、耕地面积、人口数、房屋数等的空间分布时，可以对发达地区和欠发达地区采用不同的空间关系（网格大小）来统计，并不存在从大比例尺专题图对地物（如房屋、街道、不同土地分类等）进行逐级综合到成小比例尺专题图后再作统计的必要。

纸面上的空间是有限的。当研究范围扩大后，才出现通过由大比例尺地图向小比例尺地图的综合问题。而面对海量空间的计算机网络时，研究范围的扩大并不要求先对空间数据和空间信息进行综合，缩小数据量后再进行分析。

## 2.3. 空间数据标准的发展状况

信息共享的关键是信息的标准化问题。标准化是技术性、管理性兼长的综合手段，是在生产力发展的一定阶段，根据经济、技术、管理等社会实践的需要，对重复性事物和概念所作的统一规定的过程，以建立社会秩序和获得最佳的经济社会效益。空间信息来自对于地理空间实体的抽象和描述，是对其空间特征和属性特征的内涵表达。空间信息系统又通过计算机网络技术对空间信息集进行了概念

化抽象和结构化组织管理，是一种可用于建立地理空间对象描述标准的载体。因此空间信息系统共享的标准化贯穿在空间信息采集、处理、分发、传输、应用的全过程。

地理信息共享早就引起了世界各国的广泛重视。自 1994 年以来由几十个国家的专家组成的 ISO/TC211 专业委员会，正在为全球制订 34 个国际标准，促进国际信息交换与共享而努力。这些标准由 3 方面组成，即：数据标准化(例如“空间数据交换标准”)、技术标准化(例如“GIS 软件互操作标准”)和应用标准化(例如“GIS 应用互操作标准”)。

国外地理信息标准化工作大体可分为两部分。一是以已经发布实施的信息技术(IT)标准为基础，对于空间信息共享而言，可以直接引用或者经过修编采用。二是研制地理空间数据标准，包括数据定义、数据描述、数据处理等方面的标准，是空间信息共享所特需的。

就标准的级别和内容而论，国外地理信息标准分为五个层次，即国际标准、地区标准、国家标准、地方标准、其他标准。国际地理信息标准是最高一层的标准，一般为推荐性标准，实用于全球各国。与地理信息国际标准研制有关的国际组织主要有：国际水道测量组织(IHO)、国际制图协会(ICA)、国际标准化组织下设的地理信息/地球信息业技术委员会等。与地理信息标准有关的最主要的机构是地理信息/地球信息业技术委员会，编号为 ISO/TC211。

标准体系是一定范围内的标准，按其内在联系形成的科学有机整体。ISO 作为最大的国际标准化组织之一，其下设的技术委员会 ISO/TC211 的工作范围为数字地理信息领域标准化。其主要任务是针对直接或间接与地球上位置相关的目标或现象信息制订一套标准，以便确定地理信息数据管理(包括定义和描述)、采集、处理、分析、查询、表示、以及在不同用户、不同系统、不同地方之间转换的方法、工艺和服务。该技术委员会目前正开展 34 个标准项目的研制工作，其内容主要涉及地理信息框架和参考模型、地理空间数据模型和算子、地理空间数据管理、地理空间数据服务和专用标准等，标准项目包括关系模型、地理信息术语、一致性与测试、空间模式、时间模式、数据分类方法、数据质量、空间参照系统、元数据、数据编码、影像和栅格数据、实用标准等。

目前我国在空间信息共享方面的标准制订相对滞后，当前已经出台的主要标集中在专业的代码标准，支持空间数据网络共享的标准仅有空间数据格式转换标准，为此这一方面的标准化工作任重道远。为了实现网络环境下空间数据共享及现有工作基础，还需要研究和完善空间数据元数据标准，制定地理空间信息编码规则，修改完善地球空间数据转换格式标准，修改完善基础地理信息符号标准，制定地理空间信息互操作标准等等。

#### 2.4. 大地坐标系的发展及对空间信息系统的影响

大地坐标系一直是大地测量中最基本的问题[11]。空间数据是各种测量手段的到的结果，空间信息与大地坐标密不可分。我国大地坐标系的建立经历了从参心坐标系到地心坐标系的发展过程。我国目前主要使用的坐标系有：全国统一的法定国家大地坐标系，高精度地心坐标参考框架，地方坐标系。

全国统一的法定国家大地坐标系。20 世纪 50 年代到 20 世纪 70 年代，我国以 54 坐标系为基础完成了大量的测量工作，并依据这个坐标系，采用分区平差法建立了我国天文大地网，测量了各种比例尺地形图。但由于当时条件所限，54 坐标系存在明显不足。为此国家测绘局和总参测绘局在 20 世纪 80 年代共同建立了我国 80 坐标系和新 54 坐标系。使得精度大大提高；在全国范围内，参考椭球面和大地水准面符合很好。54 坐标系与 80 坐标系均属参心坐标系。其特点是短距离精度高，密度大，使用时间长。但由于客观条件限制，参心坐标系未与地心发生联系，加之缺乏高精度的外部控制，长距离精度较低，在空间技术推广应用的今天，难以满足用户的需求。同时更不能满足中国数字地球框架的需要。目前的数字化的空间数据的一个主要来源是纸质地图数字化得来的，属于 54 坐标系与 80 坐标系，或者是城市地方坐标系。

高精度地心坐标参考框架。从 70 年代中后期开始，我国在空间大地测量方面以取得了一批重要的成果，建立了多种地心坐标系。地心坐标系的特点是精度高，直接与世界坐标系发生联系，并将随着科学技术的发展和服务目标需求而不断发生变化。但密度还远远满足不了国民经济建设的需要。

54 坐标系和 80 坐标系以及部分城市的地方坐标系是 2 维坐标系统，地心坐标系是三维系统。目前 GPS 测量的数据属于 WGS84 地心坐标系。实用中这几种坐标系并存，即 54 坐标系、80 坐标系和地心坐标系。造成我国在相当长的时间内不同的部门、不同地区、不同的需求、不同的年代所使用的坐

标系统不一致，从而给国民经济的各个应用部门、基础建设、军事应用带来极大的不便。GIS 空间数据以这些坐标系统的数据为基础，在数据共享和应用服务带来的不便是很自然的。

大地坐标系，即全球时空基准与框架，总是随着技术进步而不断精化，它的变化对传统空间数据表达的数字地图带来的麻烦是至今无法解决的问题。特别要指出的，由于 GPS 技术和整个卫星大地测量、卫星重力测量等技术的飞速发展，全球时空基准与框架不断精化，其周期越来越短，必将走向实时动态化。因此，以存贮某一坐标系下的坐标串为主要方式的空间信息系统是适应不了这种变化的，需要从地理空间数据在计算机中的表示方法来寻求非地图表示的新方法。

## 2.5. 空间索引

空间数据是包括不同时空和不同尺度数据源的集成。早期的 GIS 软件，大多采用文件系统存储空间信息，属性则存储在关系数据库中。空间信息一般采用网格索引。由于 GIS 应用的逐步推广和关系数据库技术的发展，基于关系数据库或者对象关系数据库的空间数据管理正在逐步成为 GIS 发展的潮流。因此，研究基于关系数据库技术下的空间数据组织成为当前 GIS 研究的趋势。目前许多关系数据库平台厂商推出了一系列支持空间数据的数据库管理平台，如 Oracle Spatial, Spatial Informix, DB2, Sybase Spatial Extender 等。

目前，大多数空间数据库平台厂商都提供了自有的空间索引技术，如 Oracle Spatial 就采用了四叉树和 R 树作为其空间索引。但是，不同的空间数据库平台之间的空间索引却不能通用，Oracle 空间数据库的索引机制就不能应用于 Informix 数据库中。这就给数据库应用者和数据库开发人员造成了很大不便，他们不得不面对各种不同的开发环境和开发模式，与数据库技术发展的通用化，模块化，接口标准化背道而驰。因此，建立在通用关系数据库管理系统的空间数据索引机制的研究就越来越有其价值。

国内外常见的空间索引一般是自顶向下、逐级划分空间的空间索引，比较有代表性的包括网格空间索引、四叉树索引、BSP 树、KDB 树、R 树、R+树，分别简述如下：

网格空间索引：思路简单明了，容易理解和实现。其基本思想是将研究区域用横竖线条划分大致相等和不等网格，记录每一个网格所包含的空间实体。当用户进行空间查询时，首先计算出用户查询对象所在的网格，然后再在该网格中快速查询所选空间实体，这样一来就大大地加速了空间对象的查询速度。

四叉树索引：它是在规则格网的基础上，按空间目标在格网上的分布密度不同进行一分为四的逐级划分。目标密集的区域，四叉树的格网小，目标稀疏的区域，格网大。查找空间目标时，首先查找四叉树格网包含的空间目标，加快查询速度。

BSP 树空间索引：是一种二叉树，能很好的与空间数据库中空间对象的分布情况相适应，但深度较大，对各种操作均有影响。

KDB 树空间索引：是 B 树向多维空间的一种发展。它对于多维空间中的点进行索引具有较好的动态特性，删除和增加空间点对象也方便，其缺点是不直接支持占据一定时间范围的空间对象，如二维空间中的线和面。

R 树和 R+树索引：R 树是一个与 B 树类似的动态平衡树，根据地物的最小外包矩形建立，可以直接对空间中占据一定范围的空间对象进行索引。由于 R 树结点对应的空间区域可以重叠，因此，R 树可以较容易地进行插入和删除操作；但正因为区域之间有重叠，空间索引可能要多条路径进行搜索后才能得到最后地结果，因此，其空间搜索的效率较低。

在 R+树中，结点对应的空间区域没有重叠，使得空间索引搜索的速度大大提高，但由于在插入和删除空间对象时要保证结点对应的空间区域不重叠，而使插入和删除操作的效率降低。

目前国内外现在一些大型的 GIS 软件和空间数据库系统有许多采用 R 树和 R+树建立空间索引，如 Oracle 数据库等。

## 2.6. 空间仓库技术

数据库系统作为数据管理手段，从它的诞生开始，就主要用于事务处理。随着技术的进步，人们试图让计算机担任更多的工作，而数据库技术也一直力图使自己能胜任从事务处理、批处理到分析处理的各种类型的信息处理任务。后来人们逐渐认识到，在目前的计算机处理能力上，根本无法实现这种功能，而且，事务处理和分析处理具有极不相同的性质，直接使用事务处理环境来支持决策是行不通的。要提高分析和决策的效率和有效性，分析型处理及其数据必须与操作型处理及其数据相分离。

必须把分析型数据从事务处理环境中提取出来，按照决策支持系统（DSS）处理的需要进行重新组织，建立单独的分析处理环境，数据仓库正是为了构建这种新的分析处理环境而出现的一种数据存储和组织技术。

各类 GIS 应用、决策需要从大量的信息上进行处理和分析，空间信息在这种分析处理中起作极其重要的重用。由于各种类型的空间信息的来源不同、格式不同、类型不同、时间不同，进行决策分析是非常复杂和费时的。GIS 在取得巨大发展的同时，多维信息的空间分析能力不足以及空间分析的结果不实用越来越明显。在实际应用中，使用者不仅要从 GIS 中得到详细的多维信息，也要得到概括的多维信息，这些信息往往对提高空间辅助决策分析的能力非常有用。为此有必要引入数据仓库技术。数据仓库在空间数据库上表现为空间数据仓库。

90 年代以来，以美国为首的发达国家积极开展空间信息处理、空间数据仓库和数据挖掘技术的基础理论研究和应用研究。90 年代中期以后进一步强调了空间数据仓库、空间数据联机分析和空间数据挖掘的研究，并作为一个重要的研究方向。我国在空间数据仓库、空间数据联机分析和空间数据挖掘的研究方面还比较落后。发展空间信息共享，建立空间数据基础设施，空间数据仓库、空间数据联机分析和空间数据挖掘是一项基础研究工作。

然而，目前的空间数据仓库的研究主要依托于现有的 GIS 提供的结构。目前的 GIS 已不足以描述地理要素的多维信息结构[12]，从而不能方便地进行多维信息地空间分析和多维信息地概括分析。这已严重地制约了 GIS 地发展。因此空间数据仓库技术的发展，需要有新的地理要素的多维描述结构的支持。

综上所述，从目前地理信息系统与相关技术的发展及相互集成中存在的问题，和从数字地球到数字城市及其它应用所提出的要求两大方面来看，需要人们突破几千年来传统的以地图方式表示地理空间数据和信息的框框，去探求地理空间数据与信息在网络计算机系统新的表示方法，去创造与全球开展的网络计算相适应的空间数据表示的新方法，这就是本文探索性研究的出发点。

### 3. 网络技术——Internet 的第三次浪潮

如上文所述，网络是高性能计算机、数据源、因特网三种技术的有机组合和发展，它把分布在各地的计算机连接起来，用户分享网上资源，感觉如同个人使用一台超级计算机一样。从数量上来说，网络的带宽更高，计算速度和数据处理速度大幅提高，结构体系比现有网络更有效的利用信息资源。网络的根本特征是资源共享，消除资源孤岛。

目前网络技术研究工作主要在网格计算、信息网格等方面。

网格计算（Grid Computing）通过网络连接地理上分布的各类计算机（包括机群）、数据库、各类设备和存储设备等，形成对用户相对透明的虚拟的高性能计算环境，它的应用包括分布式计算、高吞吐量计算、协同工程和数据查询等诸多功能。网格计算被定义为一个广域范围的“无缝集成和协同计算环境”。

信息网格是要利用现有的网络基础设施、协议规范、Web 和数据库技术，为用户提供一体化的智能信息平台，其目标是创建一种架构在 OS 和 Web 之上的基于 Internet 的新一代信息平台 and 软件基础设施。在这个平台上，信息的处理是分布式、协作和智能化的，用户可以通过单一入口访问所有信息。信息网格追求的最终目标是能够做到服务点播（Service On Demand）和一步到位的服务（One Click Is Enough）。信息网格的体系结构、信息表示和元信息、信息连通和一致性、安全技术等是目前信息网格研究的重点。

网格作为 Internet 的第三次浪潮，在各国引起了前所未有的关注和重视，计算机方面的世界级大公司竞相推出了网格技术基础平台和相关的协议标准。

目前，网格研究主要以美国和欧洲为首，尽管网格技术还远不如互联网和 Web 技术那么成熟，但却已有政府、部分公司和研究所进入了使用或试验阶段。其中英国政府已投资 1 亿英镑，用以研发“英国国家网格”（UK National Grid）。美国政府用于网格技术的基础研究经费则高达 5 亿美元。美国军方正规划实施一巨型网格计划，名为“全球信息网格”（Global Information Grid），预计在 2020 年完成。作为该计划的一部分，美国海军陆战队另推动一项耗资 160 亿美元、历时 8 年的项目，包括系统研发、制造、维护及升级。美国能源部的山地亚国家实验室也宣布，其“先进战略运算创新计划网格”将用于核武器研究，此外，美国国防部与欧洲能源机构等都已先后开始采用网格技

术。论及网格的应用范围，放眼国内诸多领域，如能源、交通、气象、水利、农林、教育、环保等都对高性能运算网格及信息网格的需求有迫切的要求。

产业界已经在大力推动这最后一种工作，包括惠普公司的 eSpeak 和 e-service、IBM 公司的 Web Services、微软公司的 .Net、以及太阳微系统公司的 Open Net Environment (Sun ONE)。这些公司已经在使用 XML、SOAP、UDDI 等标准接口方面达成共识。

在国内，中国科学院计算技术研究所对网格技术的研究已较为深入，其研究的项目称为织女星网格，研究工作分为三个层次，即计算网格、信息网格、知识网格。

在网格的体系标准方面，国际上有一个非常大的组织 OGSA (Open Grid Services Architecture)，制定网格服务体系标准。OGSA 已经成为了国际上公认的标准，IBM 在倡导推广。另外还有一个全球网格论坛 GGF(Global Grid Forum)，是一个独立从事研究、开发、发布和支持有关网格活动的组织。网格论坛里已经有很多单位都围绕 OGSA 来做研究和应用。

网格技术与空间信息处理结合（称空间信息网格）的研究目前还刚刚开始。空间信息网格目前主要涉及三个技术：网格计算、基于 Web 服务的信息共享技术和空间信息互操作技术。前两项属于信息基础设施建设和技术标准问题。后一问题，两个国际化标准组织 OGC 和 ISO/TC211 已经作了大量的研究工作，发布了一系列技术标准，如基于 Web 的地图服务规范 (WMS) 和基于 Web 的要素服务规范 (WFS) 等，从技术上解决了异构数据库的互操作问题。但是如果数据存在坐标系不一致、基准不一致、数据的比例尺不一致等问题，从不同的数据源获得的数据产生矛盾，即使能相互叠在一起，也会产生应用上的矛盾。因此，空间信息网格需要适合于网格技术的空间信息表示、空间数据组织等方面的技术支持。

在地球空间全球网格划分研究方面，全球地理信息多尺度表示和数据组织一直许多人关注的问题。多年来地学领域的有关专家、学者为此进行了大量的研究探索。

1992 年 Goodchild 和 Yang 就提出了全球地理信息系统一种层次数据结构[2]。其思想在后来 Dutton 的全球层次坐标网格体系被采用和发展。

1998 年，Sahr and White 讨论了离散全球网格系统 (Discrete Global Grid System: DGGS) [13]，阐述了以经纬度来划分地球网格的方法，以及以 4 面体、立方体、八面体、12 面体和 20 面体等 5 种理想的立体形状为初始状态嵌入地球体来划分全球网格的方法。

2000 年，Dutton 在 Goodchild 思想的基础上提出全球层次坐标 (Global Hierarchical Coordinates) 方法 [14]，通过对地球进行八面体的四分三角形网格 (Octahedral quaternary triangular mesh: O-QTM) 逐级划分，形成全球多级网格 (global multi-scale mesh)，QTM 使用称为 Loc8 的二进制概念来描述位置信息。Dutton 讨论了 Loc8 编码方法和结构，阐述了 QTM 层次坐标系统。按 QTM 层次划分，划分 10 级得到约 10km 分辨率级网格，划分 20 级得到约 10m 分辨率级网格，划分 25 级得到约 1m 分辨率级网格，划分 30 级得到约 1cm 分辨率级网格。

但是上述网格的划分，在技术层面上更多地属于用于解决 GIS 空间定位、空间检索机制的网格划分，对于解决当前 GIS 所使用不同坐标系统基础数据管理，这一 GIS 的先天不足没有根本性解决。为此，我们提出的‘地球空间信息多级网格’立足于上述网格技术的基础上，重点考虑网格的基准问题、网格的投影技术和网格内不同坐标系的基础数据表示、转换方法和标准体系。

地球空间信息多级网格中的“网格”概念与网格技术中“网格”是两个不同的概念，前者主要立足于对地球空间的划分、空间数据的组织以及检索等技术，而后者主要强调在广域网上整体资源的整合与利用。然而，在广域网络环境下的空间信息资源如何有效整合才能够有利与空间信息的共享与利用，目前的技术不能很好解决。本文作者试图将“网格”的两种概念加以结合，以地球空间信息多级网格的空间划分、空间数据组织与表示方法作为网格节点上空间数据组织与管理的基础，便于网格计算环境下空间信息资源的整合、共享与利用。提出的建立空间信息多级网格体系结构正是为了适应这一要求的一种大胆的尝试。

#### 4. 空间信息多级网格理论方法体系的初步设想

基于以上对现有空间信息技术的分析和网格计算时代的即将到来，空间信息走网格化道路是大势所趋。但是考虑到地球空间的自然特征和社会属性的差异性以及和经济发展的不平衡的特点，本文提出一种适合网格计算环境支持下的空间信息多级网格技术。下面简要介绍作者所提出的空间信息多级网格 (Spatial Information Multi-Grid: SIMG) 的体系结构、信息表示与数据存取、与传统空间信

息系统之间的数据转换、空间信息多级网格处理与分析的技术以及空间信息多级网格的主要应用领域。

#### 4.1. 空间信息多级网格的体系结构

空间信息多级网格的核心思想是：按不同经纬网格大小将全球、全国范围划分不同粗细层次的网格（如四层），每个层次的网格，在范围上具有上下层含盖关系。每个网格以其中心点的经纬度坐标（网格中心点）来确定其地理位置，同时记录与此网格密切相关的基本数据项（如经纬度，全球地心坐标，各类投影参数下的坐标）。落在每个网格内的地物对象（细部地物）记录与网格中心点的相对位置，以高斯坐标系或其它投影坐标系为基准。根据实际地物的密集程度确定所需要的网格尺度（分层密度），如地物稀疏的地方只需要粗网格，而地物密集的地方（如城市）则按细网格存贮空间与非空间数据。

将不同的网格层次同全国、省、地市、县等行级别建立关联，同时结合我国国家级、省市级空间信息基础设施的建设和信息网格技术，建立我国空间信息多级网格（SIMG）的体系结构，具体内容包

- 括：
- (1) 空间信息多级网格的划分。确定多级网格的层次数，各级网格的大小，不同地域网格粗细层次程度的确定原则。
  - (2) 每个网格点属性项的确定。包括应具备哪些基本属性、自然属性、社会属性、经济属性、文化属性……。
  - (3) 行政区划与空间信息多级网格对应关系的确定。这为以行政区划目标的信息统计、宏观分析与决策提供基础。
  - (4) 基于网格计算技术的空间信息多级网格结构。研究空间信息多级网格结构与网格计算技术结合，提供空间信息服务的体系与服务模式。

这四项内容应当成为标准，数据采集、更新、维护、管理和使用部门均服从这个统一的空间信息多级网格标准，以利于网格分布计算。

空间信息包括几何与属性数据按多级网格的存取与多比例尺地图空间数据库的区别在于所有数据直接按由粗和细网格组成的统一网格系进行存贮、查询、分析和应用，而不需要通过地图综合为每一级网格（每一种比例尺—）存贮一个完整的数据集。即所形成的是一个变网格的统一数据集。从而保证了数据的完整性和一致性。

#### 4.2. 空间信息多级网格的信息表示、数据存储与快速查询技术

以传统 GIS 空间数据表示、数据组织和查询技术为借鉴，研究空间信息多级网格的信息表示、数据存储与快速查询技术。具体内容包

- 括：
- (1) SIMG 数据表示。SIMG 数据同样包括几何数据和属性数据。传统的 GIS 几何数据存储用一个统一的坐标单位，而 SIMG 空间数据存储的是相对与每个网格中心点的相对坐标，这种表示方法可以保证由于参考系变化、测量技术进步而带来的数据变化时只改变网格中心点坐标，细部地物的增量坐标不用改变，同时具有数据压缩的能力。SIMG 数据表示需要研究这种方法带来的空间几何数据表示上变化，与属性数据表示的关系，数据输出表示上带来的差别。同时，研究基于 SIMG 的矢量数据与栅格数据的一体化表示。
  - (2) SIMG 数据和网格编码方法。需要研究不同层次的网格的统一编码方法，以及网格内部空间目标的编码。编码方法要求能够做到将不同比例尺层次的空间信息以一个一体化的数据库进行统一存储与管理。
  - (3) SIMG 数据组织。由于 SIMG 方法以网格中心来表示有关的信息，这为信息的统计、分析与宏观决策带来了很大的好处。按 SIMG 网格组织，与网格计算相结合，按不同层次的网格建立分布式的空间信息服务，在进行宏观经济统计、分析与决策时可以转化为一般的属性分析过程，做到不需要空间几何运算参与。为了达到这个目标需要研究基于 SIMG 的数据组织方法。
  - (4) 基于 SIMG 数据存储。基于 SIMG 结构，数据包括网格层次数据和网格内部数据，这两部分如何存储，需要深入研究和对比分析。
  - (5) SIMG 快速查询技术。按空间信息多级网格的划分思想，空间目标存在固有的两级索引，一是网格间索引，二是网格内部索引。可以分析目前的索引方法，比较选出或提出基于 SIMG 的索引方法，实现空间信息的快速查询。



#### (6) 基于 SIMG 的数据更新方法研究。

只有解决了这里所列出的技术关键点，空间信息多级网格才有可能成为现有数字地图和空间数据库的补充或替代产品得到真正的应用。

#### 4.3. SIMG 与各种比例尺地理空间信息系统之间数据的自动转换

SIMG 是一种空间信息表示与组织的新思想和新方法，以这种结构建立的空间数据库，其数据最初必然来自现有的空间数据库，当然，SIMG 的表示适合于直接测量数据的存储。主要研究内容包括：

- (1) SIMG 与各种比例尺地理空间信息系统之间的自动转换算法。
- (2) 转换尺度与信息损失的研究。
- (3) SIMG 与 4D 产品的转换工具。

解决了这个问题，就可以尽可能地借助已有的空间数据产品来自动建立空间信息多级网格，做到对已有空间数据成果的继承。

#### 4.4. 空间信息多级网格处理与分析技术

基于 SIMG 处理分为两个层次，一是网格层次，一是网格内部。网格内部的处理与分析，可以在传统 GIS 空间分析的基础上进行，同时结合网格计算技术，实现空间网格信息结点内的处理与分析服务。网格层次上的处理与分析需要空间网格信息结点上之间的协同，同时考虑 SIMG 的特点。主要内容包括：

- (1) 基于 SIMG 结点内部的处理与分析技术。
- (2) 基于 SIMG 的处理与分析模型与服务模型。
- (3) 空间信息多级网格处理与分析基本算法的研究与实现。
- (4) SIMG 与空间数据仓库、空间数据立方体结合，进行联机数据分析技术研究。

这里提出的空间信息多级网格处理和分析的技术关键到它能否有效地融合到网格计算的环境中，来满足经济建设、国防建设和面向全社会服务的需要。

#### 4.5. 空间信息多级网格在国家和省市宏观决策中的应用研究

空间信息多级网格研究的主要目标是空间信息的共享与服务。以网格中心点为数据附属体，通过网格内部空间数据处理与分析，提供以网格为主题的数据，这些数据构成基于多级网格信息的基础，在 SIMG 软硬件环境的支持下，为国家和省市宏观决策提供空间信息服务。主要内容包括：

- (1) 基于 SIMG 的、面向国家和省市宏观决策的主题信息组织结构研究。研究主题可以面向教育、人口、资源环境、交通等领域。
- (2) 以人口普查为例，研究空间信息多级网格在国家和省市宏观决策中的应用。

空间信息多级网格作为一种新的空间信息在计算机中的表示方法，其最主要而且最有效的应用应当是在国家和省市自然、经济、社会发展信息的及时获取、分析和在宏观决策中的应用。由于其数据结构与传统空间数据库的差别，在这类应用中空间信息多级网格将发挥其独特的优越性。

#### 5. 结束语

与几千年使用地图的历史相比较，电子计算机问世才五十年的历史，把空间信息放到电子计算机中去建立地理信息系统的历史就更短了。

地球是一个复杂的球体，在数学上用椭球体描述。地球表层是人们赖以生存和繁衍的空间，它由岩石圈、水圈、大气圈和生物圈四大圈层组成。自然界的演变，人类社会的活动使其在时空上不断演化。

人类社会的技术进步已把我们推向信息时代，电子计算机、网络通信、网格计算为我们创造越来越好的 Cyber (赛伯) 空间。

本文从这些客观现实出发，探讨了作为地理空间信息在网络计算机环境下表示的一种新方法的空间信息多级网格，它的概念、技术关键和应用前景。希望能使大家放思路去大胆求索。文中不妥之处，欢迎批评指正。

#### 参考文献

- [1] Li, D. and Gong, J.Y.. A unified data structure based on linear quadtrees .ZPF, 4/92.
- [2] Goodchild, M.F. and Yang, Shiren. A hierarchical data structure for global geographic information systems. Computer Graphics, Vision and Image Processing, 1992, 54(1), 31-44.

- [ 3 ] Ian Foster, CarlKesse Iman, Steven Tuecke. The anatomy of the grid: Enabling scalable virtual organizations. International Journal Supercomputer Applications, 2001,15(3):200~222.
- [ 4 ] 李国杰. 信息服务网格——第三代 Internet. <http://moilaetitia.topcities.com/grid/liquojie.htm>, 2001.9.
- [ 5 ] 龚健雅. 地理信息系统基础. 科学出版社, 2001.
- [ 6 ] 李爱勤. 无缝空间数据组织及多比例尺表达和处理研究. 武汉大学博士学位论文, 2001.5,2-3.
- [ 7 ] Bundy, G. L., Jones, C. B. and Furse, E.. Holistic Generalization of Large-Scale Cartographic Data. In: J. C. Müller, J. P. Lagrange and R. Weibel (eds.). GIS and Generalization: Methodology and practice. London: Taylor & Francis. 1995, 106-120.
- [ 8 ] Jones, C. B., Bundy, G. L. and Ware, J. M.. Map Generalization with a Triangulated Data Structure. In: Cartography and Geographic Information System, 1995, 22(4): 317-331.
- [ 9 ] McMaster, R. B. and Veregin H.. Multiple Representations of Spatial Data. In: UofM UCGIS Research Priority Nomination #4. <http://www.ucgis.org>. The University Consortium for Geographic Information Science (UCGIS), 1996.
- [ 10 ] Lee, D.. Making Databases Support Map Generalization. In: GIS/LIS ' 96. 1: 467-480.
- [ 11 ] 顾旦生, 张莉, 程鹏飞, 王权等. 我国大地坐标系发展目标. 测绘通报, 2003.3,1-4.
- [ 12 ] 邹逸江. 空间数据立方体的研究. 武汉大学博士学位论文, 2002,4-5.
- [ 13 ] Kevin Sahr and Denis White. Discrete global grid systems, Computing Science and Statistics. 30 ed. S. Weisberg, Interface Foundation of North America Inc., Fartax Station, VA., 1998.
- [ 14 ] Geoffrey Dutton. Universal geo-spatial data exchange via global hierarchical coordinates. International Conference on Discrete Global Grids, 3/2000, 1-15.